# BALIKESİR UNIVERSITY

# NFE EJSME

## Necatibey Faculty of Education Electronic Journal of Science and Mathematics Education

December 2024

Volume 18 Issue 2

efmed@balikesir.edu.tr e-ISSN: 1307- 6086 https://dergipark.org.tr/en/pub/balikesirnef

# BALIKESİR UNIVERSITY

## Necatibey Faculty of Education Electronic Journal of Science and Mathematics Education

Balıkesir University Çağış Campus 10145 Balıkesir
Tel: +90 (266) 612 14 00-08
Fax: +90 (266) 612 14 17
e-mail: efmed@balikesir.edu.tr
https://dergipark.org.tr/en/pub/balikesirnef

Necatibey Faculty of Education
Electronic Journal of Science and Mathematics Education

ISSN: 1307-6086

Necatibey Faculty of Education Electronic Journal of Science and Mathematics Education is an international on-line, refereed science and mathematics education journal that is published at least two issues in a year. NFE EJSME is firmly established as the authoritative voice in the world of science and mathematics education. It bridges the gap between research and practice, providing information, ideas and opinions. It serves as a medium for the publication of definitive research findings. Special emphasis is placed on applicable research relevant to educational practice, guided by educational realities in systems, schools, colleges and universities. The journal comprises peer-reviewed general articles, papers on innovations and developments, research reports.

All research articles in this journal have undergone rigorous peer review, based on initial editor screening and anonymized refereeing by at least two anonymous referees. NFE EJSME is an open access journal which means all content freely available without any charge. We support the rights of users to "read, download, copy, distribute, print, search, or link to the full texts of these articles".

Necatibey Faculty of Education
Electronic Journal of Science and Mathematics Education

# EDITORIAL BOARD

**Necatibey Faculty of Education**
**Electronic Journal of Science and Mathematics Education**

DECEMBER 2024 VOLUME 18 ISSUE 2

CONTENTS

Necatibey Faculty of Education
Electronic Journal of Science and Mathematics Education

# EDITOR'S NOTE

Dear Readers,

After an intensive working process, as the editorial board of Necatibey Faculty of Education Electronic Journal of Science and Mathematics Education, we are very pleased to publish the 2nd issue of the 18th volume of our journal with our authors and referees. In this issue, there are 12 research articles on different topics. We hope that these articles will be read with interest make valuable contributions to the field and shed light on further research. We would like to thank our editorial team for their efforts in the publication of this issue, our referees who contribute to the quality of our journal with their meticulous article evaluations, and our authors who contribute to our journal with their qualified and original articles. We thank you for all the support you have given to our journal so far and will give in the future.

Editor

Dr. Mustafa ÇORAMIK

Necatibey Faculty of Education Electronic Journal of Science and Mathematics Education
Vol. 18, Issue 2, December 2024, pp. 219-247. ISSN: 1307-6086

Research Article

# Examining Game Designs Compatible with the Planning, Implementation, Evaluation Model in the Context of Science Education

**Sibel YUMAK[1], Kübra ÖÇSOY[2], Saliha ÖZTÜRK[3], Elif BENZER[4]**

[1] Marmara University, Institute of Educational Sciences, Türkiye, sibelygt@gmail.com, http://orcid.org/0009-0005-8422-577X

[2] Marmara University, Institute of Educational Sciences, Türkiye, kubra.ocsoy@gmail.com http://orcid.org/0009-0006-4995-6478

[3] Marmara University, Institute of Educational Sciences, Türkiye, salihasaadetozturk@gmail.com, http://orcid.org/0009-0003-4806-3529

[4] Marmara University, Atatürk Faculty of Education, Türkiye, epehlivanlar@gmail.com, http://orcid.org/0000-0002-2518-768X

Received : 18.11.2023 Accepted : 24.07.2024

Doi: https://doi.org/10.17522/balikesirnef.1392825

---

*Abstract* – This study aims to examine the general features, quality, PIE (Planning, Implementation, Evaluation) model compliance, and environmental literacy of games created by pre-service science teachers using the PIE methodology. The research participants were pre-service science teachers who were enrolled in the elective course "Science, Sustainability and Environment (SSE)" as third-grade undergraduate students. The pre-service teachers were instructed to create a game using the PIE model while taking into account the environmental learning outcomes listed in the Ministry of National Education's (MoNE) 2018 Science Curriculum. Three researchers assessed each of 13 designs independently, and reliability was determined using Miles and Huberman's formula (1994). The study's findings showed that the majority of the games were card games, and they were authentic. It was determined that the games met the PIE model's requirements for quality and compatibility. It was also established that the pre-service teachers primarily incorporated environmental awareness and subsequently knowledge in their games as part of the environmental literacy components.

*Keywords*: Educational games, science education, the PIE model, environmental literacy.

------------

Corresponding author: Sibel YUMAK, sibelygt@gmail.com

## Introduction

Since ancient times, one of the fundamental forms of entertainment for kids has been games. Researchers studying archaeology have found drawings and reliefs that show children were playing games even in antiquity. Despite the fact that games have been around for as long as humans, their educational implications date back to the 19th century (Aymen Peker, 2018). Games are now seen as important tools for education rather than just a way to enjoy time, thanks to a change in public perception (Gee, 2001; Kirriemuir & Mcfarlane, 2007). Games that are designed to fulfill a specific educational objective while considering the number of players, skill level of players, duration and tools and equipment to be used are referred to as educational games (Özen et al., 2019).

Every stage of a person's development can benefit greatly from playing educational games. A review of the literature on educational games led researchers to the conclusion that, from pre-school through high school, games have aided in the realization of some educational objectives (Durdut, 2016; Herodotou, 2018; Şentürk, 2020). Educational games give students a chance to participate actively in the learning process, enhance their critical thinking abilities, make learning more fun, and facilitate peer learning (Beker Baş & Karamustafaoğlu, 2020; Kocadere et al., 2019; Serdaroğlu & Güneş, 2019). In addition to these advantages, using educational games in science classes helps students develop positive attitudes toward science classes by helping them to hypostatize abstract concepts (Korkmaz, 2018; Özkan et al., 2020).

Examining science education curriculum reveals that curriculum has specific objectives, including teaching environmental science, helping individuals understand the world, fostering interaction, and increasing environmental consciousness (Ministry of National Education (MoNE), 2018). The purpose of environmental education is to raise and educate people who can solve environmental issues, safeguard the environment, and are sensitive to it (Mutlu Karanfil, 2022). Encouraging individuals to act responsibly when it comes to the environment is the primary goal of environmental education (Fidan Yazgan, 2023). One could argue that the objectives of environmental education and the previously stated science education objectives are similar. Environmental literacy is brought about when this context is assessed, and the objectives are turned into actions or observable behaviors (Bilim, 2012). Environmental education aims to create environmentally literate people by preparing students for scenarios they may face in the real world. The learning environments must be suitable to produce solutions to problems that are not encountered, enabling the student to participate actively (Külegel, 2020). Educational games can be considered one of the techniques involving students’ active participation in the process. The study by Kırmızıyüz et al., (2021) concluded that educational games positively influenced the students' perspective on environmental issues, and that more environmental concepts could be learned through games. Studies on environmental units with games have revealed notable changes in the attitudes of students toward the environment (Kefeli et al., 2018).

When considering the research, it can be concluded that educational games are likely to be efficacious in both science courses and environmental topics. Specific guidelines must be adhered to in order to enhance the efficacy of educational games pertaining to the topic and lesson. When creating educational games, it is important to take into account the topic or content of the lesson, the age of the students, their levels of cognitive and emotional development, their preferred learning styles, and any individual differences (Beker Baş & Karamustafaoğlu, 2020). Games ought to be designed and implemented within a particular framework, and this procedure should be undertaken attentively (Karamustafaoğlu & Kılıç, 2020). The following features should be considered in the designed game (Aymen Peker, 2018):

- It should be learning-oriented and clear and meet students’ needs.
- It must be able to adapt to different situations.
- It should be performed in an attention-grabbing manner and in accordance with the intended objective.
- The suitability of the content should align with the specific context and qualities of the target audience, such as their age and educational level, time and physical conditions of the environment.

Research conducted with the pre-service science teachers has demonstrated that educational games enhanced the retention of learning, and that the games developed were qualified and educational. Nevertheless, the pre-service teachers expressed difficulties in finding and developing games during the game design procedure and expressed a desire for additional classes on how to effectively incorporate them into their teaching (Önen et al., 2012; Seçkin Kapucu & Çağlak, 2018). Research has shown that digital games, including those with educational purposes, enhanced the quality of the learning process, and participants expressed their willingness to create digital games using software (Akgül & Kılıç, 2020). Consequently, research has demonstrated that games can serve as an influential instrument in science education. Games facilitate students' comprehension of scientific subjects, foster their curiosity, and bolster their inclination towards exploration and discovery. The game design components must be meticulously and clearly delineated. There must be a harmonious equilibrium between amusement and instruction. Research conducted with the pre-service science teachers also revealed that participants believed that games could be utilized not only to enhance the enjoyment of lessons, but also to facilitate the achievement of learning objectives. According to the educational objectives aimed, educational games can be designed in different ways.

The PIE (Planning, Implementation, Evaluation) model was employed in this study as an alternative framework for assessing pre-service teachers' game designs. The model, developed by Jacobson et al. (2006), is a customized game design model specifically designed for environmental purposes. The proposed framework outlines a three-step design process for creating a game to protect the environment. The effectiveness of these three steps is determined by the extent to which the questions posed at each stage are answered and how they are answered. The stage of planning refers to the part when game is planned and organized. To ensure optimal performance in planning, it is imperative to address these inquiries: “What are the intended learning outcomes for the participants?”; “What is the objective you aim to achieve?”; “What type of game are you seeking?”; “Will there be a narrative incorporated?”; “Will the game be conducted in an outdoor space or inside a classroom?”; “Which gamification elements will be used?”; “Are there different difficulty levels available?”; “Will there be any rewards? What will the rewards be?”; “Which material is required?”; “Do you have these materials?”; “If you do not have these materials, do you have to provide them?”; “How can you provide them?”; “Did you conduct any pre-implementation testing on the game?”; “Does the test result need to be modified?”. The implementation phase refers to the stage in which the game is actively played. The question of whether the game can be played without any issues is significant. In the stage of evaluation, the information acquired in the game can be evaluated by posing questions like “What have you just experienced?”, “What does this mean?”, “Now what can you do with this knowledge?”, and the game can be evaluated through the questions “Where did you get the most trouble in the game?”, “What were the easy parts?”, “Which part of the game would you like to change?” (Jacobson et al., 2006).

The significance of PIE lies in its ability to engage students actively, enhance their learning experience, optimize learning outcomes, analyze performance-product-based challenges, and include the steps of planning, implementation and evaluation. PIE is a facilitating method that helps solve problems systematically, therefore providing a clear plan and suggestions for situations that arise in education. It also helps instructors by guiding them through the steps of planning, implementation, and evaluation, which helps to minimize unexpected issues (Keleş et al., 2016). Developing teaching design model aims facilitate the selection of effective designs by comparing the designs based on the educator's alignment with the intended objective and the designer's knowledge of the design's origin (Andrews & Goodson, 1980). The main objective of this study is to analyze the game designs of pre-service science teachers regarding environmental issues, using the PIE model, and evaluate those designs in terms of game, model, and environmental literacy.

In this study, the PIE model was chosen due to its compatibility with the integration of various multimedia tools and technology (McPheeters, 2009; cited in Keleş et al., 2016) for the development of products both inside and outside the classroom (Gustafson & Branch, 2007; cited by. Keleş et al., 2016). When compared to other designing tools, one of the reasons the PIE model was selected for this study, was its suitability for basic level education, because pre-service teachers were expected to design games for middle school students. Also, this model promotes constructivist learning, pre-service teachers in the role of teachers focus on coordinating learning as guides of teaching rather than simply providing information. Hence, they meet their students with a great experience after graduation (Keleş et al., 2016). This study is significant because it provides a different perspective on the use of games in science education since there are very few studies about PIE model and its implementation in Türkiye. Hopefully, this study can inspire educators to use PIE model in environmental education.

The research questions of the study are as follows:

1. What are the general features of the games designed by the pre-service science teachers?
2. What are the qualities of the games designed by pre-service science teachers?
3. How do pre-service science teachers utilize the phases of PIE when designing games centered on environmental issues?
4. Which components of environmental literacy do pre-service science teachers employ when creating games based on the PIE model?

## Method

### Research Design

In this study which analyzed the game designs of pre-service science teachers focusing on environmental issues using the PIE model, a case study approach within the qualitative research methodology was applied. Case studies involve in-depth examination of one or more situations, with the goal of establishing associations with various factors (such as medium, individuals, events, processes, etc.) and producing outcomes (Yıldırım & Şimşek, 2021). In this study focusing on the game design process undertaken by pre-service science teachers, using the PIE model and the environmental literacy in those designs constituted the case of which sub-dimensions were being examined.

### Participants

The study was conducted during the spring period of 2018-2019, involving 56 pre-service science teachers, 8 male and 48 female, enrolled in the Science Teaching Bachelor's Program. These candidates were in the third year of the program. The participants had taken courses about environmental education and games in science education before this study. Yet, they were not familiar with PIE model. Since the participants were already taking courses together, convenience sampling method was used in this study. This sampling method involves selecting participants who are readily accessible to the researcher, meaning that there is a pre-existing sample (Şimşek, 2018). The data source for the study consisted of 56 pre-service teachers, divided into 13 groups. Groups were sequentially named as G1, G2, ... and G13.

### Implementation Process

The implementation process was delineated into two fields examined in the study. These were the environmental content and literacy and game design following the PIE model. The environmental concepts of the study were taught by one of the researchers and the lecturer of the class in a 3-hour Environmental Science class attended by all the pre-service teachers. The concepts about PIE model and game design were taught by the same lecturer during a three-hour elective class Science, Sustainability, and Environment (SSE) that all participants attended. The concepts of science, sustainability, and the environment were defined in the first weeks of the course and the concepts were subsequently deliberated with the students in a questioning environment, elucidating their interrelation using everyday life illustrations. The primary scope of the study, PIE model was conducted during the eighth week of the course and spread out over a period of two weeks. Upon completion of the class, the pre-service teachers were expected to develop a game design based on the PIE model, for a specific subject or learning outcome from the MoNE 2018 science curriculum, and subsequently submit a report on it. In these reports, pre-service science teachers were expected to explain how they followed the steps of PIE model while designing an educational game. In the planning part the problem was introduced, the objective of the game, the level and prior experiences of the participants, and the limitations were described. In the implementation part, pre-service teachers were expected to describe possible changes after a pilot study and effectiveness of the planning part. In the evaluation part pre-service teachers were expected to describe the product and make suggestions to evaluate whether the method works. At the final part of the reports, participants included the game designs that they developed for teaching environmental concepts.

The pre-service teachers created this design by organizing themselves into groups consisting of their own acquaintances, with group sizes ranging from three to six individuals. It was thought that allowing pre-service teachers to be in a group with their preferred companions and to create their designs centered on an environmental theme of their personal preference would enable them to be motivated and self-assured. All pre-service teachers submitted reports as a compulsory part of the course and all the participants volunteered to have their reports used in this study. Given that the instructor had a minimum of five years of teaching experience, this study did not involve any pilot practice. Instead, the researchers relied on their own expertise.

## Data Collection Tools

The data were assessed utilizing a measurement instrument devised based on four themes. These themes were derived from the general features and qualities of the games, their compatibility with the PIE model, and the environmental literacy components they produced. Extensive literature research was conducted for each specific field while preparing the measurement instrument (Akcanca & Sömen, 2018; Freitas & Oliver, 2006; Jacobson et al. 2006; Önen et al. 2012). The examination focused on the components of environmental literacy namely awareness, knowledge, attitude, skill, and behavior. During the analysis of these game components in the designed games, the definitions of the components and the corresponding keywords derived from those definitions were utilized. Environmental literacy components included the following definitions (Alınmaz, 2023; Altınöz, 2010; Benzer, 2010; Fang et al., 2022):

1) *Awareness* refers to a broad understanding or perception of environmental issues.
2) *Information* refers to the knowledge of environmental concepts, such as ecology; this includes understanding ecological information, environmental problems, and environmental issues.
3) *Attitude* refers to an individual's capacity for understanding and showing concern for the environment, while also displaying a positive outlook and being responsive to environmental issues (such as anxiety, fear, and motivation).
4) *Skill* refers to the capacity to effectively utilize one's acquired knowledge and attitudes to address environmental issues. (This includes analyzing, generating solutions, and making decisions.)
5) *Behavior* refers to the manifestation of one's environmental knowledge, attitude, and skill, which leads to active engagement in environmental matters and concerns.

Following the completion of the research, the researchers convened to develop the initial version of the measuring instrument. The initial version of the measuring instrument comprised two primary components. The initial version comprised items pertaining to the game's quality and the steps of PIE model. To ensure the validity of items, researchers consulted both a field and a language expert. The measuring instrument was divided into four sections based on the feedback received, and adjustments were implemented in the part of scoring. Using this rubric, an evaluation was conducted on three games designed by the pre-service teachers. Following the pilot implementation, experts reached a consensus to revise certain items and omit some other items, resulting in the finalization of the measuring instrument. The final version of the rubric consisted of 14 items that described the general features of the games, 11 items that outlined their qualities, 13 items that assessed their compatibility with the PIE model, and 5 items that evaluated their environmental literacy components.

**Data Analysis**

For the first part, in the evaluation of the general features of the games, the type and authenticity of the game, and number and grade level of players and the frequency of using game elements included were expressed. In the second part, during the evaluation of the games' qualities, three researchers classified the games into four categories: "0" (no data), "1" insufficient (data available but design not suitable), "2" partially sufficient (design contains appropriate data but not enough), and "3" sufficient (design has data that fully corresponds to the item). The frequencies of each score value for the items included in the measurement instrument were also computed by three researchers. The maximum attainable score in this particular section was 33. Excerpts from sample designs that received the maximum score were shown in Table 1.

**Table 1** Explanation that Received the Maximum Score on the Game Quality

| Game | | 3 point-samples |
|---|---|---|
| Item 1 | Learning outcomes | "Our activity consisted of the group game designed in order to make students comprehend the learning outcomes stated in the unit of 8th grade Science Education titled Cycles of the Matter and Environmental Problems, namely 'Students will be able to discuss the causes of global climate change and its potential outcomes.' ." (G11) |
| Item 2 | Objective | "In our game design, it is crucial that participants learn through enjoyable experiences. Our game aims to enable participants to be individuals who possess a strong awareness of environmental issues and a comprehensive understanding of the potential consequences of global climate change resulting from these issues." (G11) |
| Item 3 | Rules | "Rule**:** Prohibited words written on the cards cannot be uttered.<br>No portion of the words written on the card may be used as a clue.<br>Sign language cannot be used. Songs cannot be explained through mere whispers of the melody or the recitation of lyrics.<br>You cannot explain a word by articulating words in any foreign language except Turkish." (G8) |
| Item 4 | Difficulty | The game's level of difficulty was assessed by analyzing the students' prior learning and the consistence between the learning outcomes and objectives targeted by the game.<br>"Prior learning;<br>At the fifth grade, students are presumed to have attained the knowledge and skills related to the learning outcomes in the unit of the Human and Environment.<br>S.5.6.2.1. Students will be able to recognize the significance of the interaction between humans and the environment, denote the detrimental impacts of environmental pollution on human well-being.<br>Learning outcomes targeted by the game;<br>S.8.6.3.3. Students will be able to discuss the causes and potential consequences of global climate change." (G2)<br>Since the designed game was designed for the 8th grades, prior learning and targeted learning outcomes are coherent. |
| Item 5 | Time management | "One group is assigned the task of monitoring the passage of time, while a member from another group is provided with a set of cards at the beginning of the time period. The individual's own group is then instructed to explain the words written on the cards within duration of 1 minute, without using any forbidden words." (G8) |

**Table 1** (continued)

| Game | | 3 point-samples |
|---|---|---|
| Item 6 | Instructing | "The objective is to teach students the determinants that impact global warming, the greenhouse effect, and the potential consequences of environmental issues on the future of the Earth and human existence." (G11) |
| Item 7 | Reinforcing | "The product of the game designed for $7^{th}$ grade students is defined as students' having a thorough understanding of which waste materials can be recycled. The game's deck comprises various waste materials, and students are required to differentiate between them. The consistency between the aimed learning outcomes and game concepts has been acknowledged as a means of reinforcement." (G4) |
| Item 8 | Evaluation | "Following the game, the teacher can understand whether objectives were realized or not by examining the quantity of stickers collected by students and administering a quiz." (G12) |
| Item 9 | Social interaction | "Related to skills, the student develops proficiencies in areas such as communication, self-expression, collaboration, and time management." (G8) |
| Item 10 | Competing | "A mysterious bicycle journey which was set on with 24 golden coins... During this expedition, we both have the opportunity to get to know the cities and acquire new information. As we learn, we embark on adventurous journeys. If we fail to do something for the environment, these adventurous paths will be our nightmares. Furthermore, we have the opportunity to engage in commerce along those paths. However, what matters more is not the money but the points. Let's see who will rack up the most!" (G10) |
| Item 11 | Active learning | "The game helped students to be active and remember previous learning." (G12) |

The evaluation of the compatibility of the game designs with the PIE model for the 13 items in the measurement instrument was similar to the evaluation made in the qualities of the game. For "0," the PIE model was never taken into account; "1" was graded as insufficient enough to the PIE model, "2" was only partially sufficient, and "3" was graded as sufficient to the PIE model. For each item, the analyses were tabulated to provide frequency. Maximum score was 39. Excerpts from the designs are shown in Table 2 as examples of the evaluation.

**Table 2** Explanations with a Maximum Score Based on the PIE Model

| Game | | 3 point-samples |
|---|---|---|
| Item 1 | Issue/problem handled | "Human-environment relation. Positive and negative impacts of human beings on environment. Harms caused by human beings on environment (Consciously or accidentally)" (G1) |
| Item 2 | Objective and purpose | "To increase sensitivity to environmental problems caused by human activities. To develop the ability to protect and improve the environment. To question recycling in terms of resource efficiency, environmental pollution and economic development." (G4) |
| Item 3 | Levels/features of participants | "They should answer the questions sincerely. They should be 5th grade students. Participants are interested in environment and environmental problems." (G1) |
| Item 4 | Prior experiences of participants | "They have information about the beneficial or harmful situations in the human-environment interaction. They have the required information and skill for the solutions to environmental problems." (G4) |
| Item 5 | Change | "Students can distinguish recyclable materials from nonrecyclable materials in the cards they are given during the game. They plan how to recover the materials written in their cards." (G4) |
| Item 6 | Planning process | "For the factors to be included in the game like the resources, types of plants, types of animals, any elements students demand can be included in the game. They can be allowed to include their own city names. Their expectations and demands can be asked before playing." (G2) |
| Item 7 | Limitations | "There may be troubles related to the accuracy of the answers given by students during the activity. Students may not be honest about their answers just because they want to win the game. It may be hard to play the game in a crowded classroom. The number of the cards may be increased based on the class size." (G1) |
| Item 8 | The impact of instructions on knowledge, attitude and behavior | "When feedback related to the mistakes of students are given via red dots until reaching the best world among four different ones, instructions may have much influence upon changing the negative attitudes of students. The examples given in the orientation parts of the game include attitudes and behaviors that are highly applicable in everyday life (e.g. greengrocer, market). So, the fact that students can demonstrate positive behavior change can be of great importance in terms of accelerating." (G7) |

**Table 2** (continued)

| | | |
|---|---|---|
| Item 9 | Implementation-pilot | "Due to the game applied, students acquire detailed information on air pollution, which is an important form of environmental pollution.<br>They learn the results of the impact of air pollution on earth.<br>They utilize the behaviors and attitudes acquired in the game. (G5) |
| Item 10 | Finance | "It is financially suitable to prepare the game. Students can prepare this game using their own stationery materials." (G5) |
| Item 11 | Working of evaluation method | "Before applying the method, a survey which aims to measure knowledge, attitude and behavior related to the topic is implemented.<br>After the method is tried, the same survey is implemented again. Comparing the first and the latter, the change is examined." (G6) |
| Item 12 | The output of the program | "Students can discern the factors that cause environmental problems, think about them, interpret them; their opinions change, and they are motivated.<br>And all those factors make students more sensitive to the environment, and their behaviors change." (G6) |
| Item 13 | Success indicators | "Yes, we make assessments via surveys implemented.<br>Also, we can understand which component of the environment challenge students most by examining at what level students often lose the game." (G6) |

To see which components of environmental literacy were included in designed games, the evaluation of the environmental literacy components in game designs was conducted independently by three researchers. Among the environmental components, the following ones were included in the scope of the study: awareness, knowledge, attitude, skill and behavior. Table 3 presents excerpts from sample designs that received the maximum score for environmental literacy.

**Table 3** Explanations with a Maximum Score Related to Environmental Literacy

| Environmental literacy component | 3 point-samples |
|---|---|
| Environmental awareness | “While evaluating the designed games regarding the environmental awareness, the inclusion or exclusion of environmental issues in games has been examined. Environmental concerns were incorporated into the majority of the games, with the intention of providing changes in the students related to the topics on environment. Ina game of which content was stated, the following explanation was made: "acquiring knowledge about household waste, its recovery and recycling, and cultivating appropriate behavior".” (G9) |
| Environmental knowledge | “To distinguish between recyclable and non-recyclable waste, accurately categorize recycled waste, and obtain information about the advantages of recycling.” (G13) |
| Environmental attitude | “The game consists of a maze with two possible exits. In the pathways of the maze, items related to the environmental attitude and behavior are included. These items are annotated with short marks such as 'yes-no' or 'do-do-not'. Via the marks, the game concludes with two possible outcomes: one leads to a nice and clean environment, while the other leads to an environment that is contaminated and inhospitable.” (G7) |
| Environmental skill | “The game, which is played by a minimum of two groups, and improving the capacity to make rapid decisions and categorize by engaging in competitive time-based challenges.” (G13) |
| Environmental behavior | “The purpose of the designed game will be to design the most environmentalist city ever, in other words, Eco-City. The virtual tabs will contain many different resources. There will be many energy options, from wind turbines, for example, to nuclear power plants. There will be a recycling center, regular storage facilities. The city will be founded by the students' choices. The student will pay attention to the characteristics that are environmental when building a city, in other words, they will display a behavior. This game is thought to contribute to environmental behavior thanks to gaining points and levels.” (G2). |

### *Reliability*

The games were assessed by three researchers independently. Below are the specifics of the data that underwent descriptive analysis using closed encoding. As a result of the conducted analyses, the consistency among researchers was examined using the matching formula proposed by Miles and Huberman (1994). For the graded parts of the rubric sections (the features of the game, components of PIE and environmental literacy), consistency of researchers were examined one by one. The researchers' consistency was measured at 0.80 for quality, the PIE components at 0.77, and the environmental literacy at 0.75. The calculated consistency values can be considered suitable in relation to the research's reliability (Miles & Huberman, 1994).

## Findings

To answer four research questions of this study, the reports, which were filled by participants and included designed games, were analyzed. The results were organized into four research questions.

### General Features of the Games

The first research question of this study was "What are the general features of the games designed by pre-service science teachers?". The general features included the type of game, game authenticity, number of players, game level and game components. The findings regarding to the first research questions are presented in Table 4.

**Table 4** General Features of the Games Designed by the Pre-service Teachers

| Type of game | f | Game authenticity | f | Number of players | f | Game level | f | Game components | f |
|---|---|---|---|---|---|---|---|---|---|
| Card game | 5 | Authentic | 8 | Individual | 4 | Grade 4 | 1 | Goal setting | 13 |
| Board game | 3 | Adaptation | 5 | Group | 7 | Grade 5 | 4 | Re-playing | 13 |
| Digital game | 3 | | | Both | 2 | Grade 7 | 3 | Giving feedback | 13 |
| Other | 2 | | | | | Grade 8 | 3 | Setting rules | 12 |
| | | | | | | Not specified | 2 | Interesting | 12 |
| | | | | | | | | Reward | 8 |
| | | | | | | | | Challenge | 6 |
| | | | | | | | | Time | 3 |
| | | | | | | | | Level | 3 |
| | | | | | | | | Narration | 1 |
| Total | 13 | | 13 | | 13 | | 13 | | 84 |

f: frequency

The data from Table 4 shows that card games (38.4%) were the most used ones among the games designed by the pre-service science teachers. The card games were followed by board games (23.1%), digital games (23.1%), and other games (15.4%). To assess the authenticity of the games, they were categorized into two groups: authentic games and adapted games. Out of the 13 games developed, 8 games (61.5%) were identified as authentic, while the remaining 5 games (38.5%) were classified as adaptations. The games were analyzed in three groups based on the number of players; namely individual games, group games, and games that allowed both individual and group playing. Out of 13 games, 4 games

(30.8%) were played individually, 7 games (53.8%) were played as a group, and 2 of them (15.4%) could be played either individually or in groups. When grade levels of the designed games were examined, the games were mostly designed for Grade 5 (30.7%), followed by Grade 7 (23.1%) and 8 (23.1%). 2 games (15.4%) did not have a level, and 1 game (7.7%) was specifically created for students in the fourth grade. Upon evaluating the games created by the pre-service teachers in terms of game components, it was found that all of the 13 games included objectives and gave feedback and were designed to facilitate re-playability. Furthermore, the rules were explained properly in 12 games (92.3%), and 12 games (92.3%) were designed in an interesting way. 8 of the games (61.5%) specified the reward, while 6 games (46.1%) incorporated the component of challenge. The games had minimal references to time and level, with only 3 games (23.1%) providing explicit specifications for both. Only one of the games (7.7%) included a narrative component.

## Qualities of the Game

To seek an answer to the sub-problem as “What are qualities of the games designed by pre-service science teachers?’’, the game's qualitative features have been examined. Findings are presented in Table 5.

**Table 5** Findings Related to the Qualities of the Designed Games

| Items | None (0) f | Insufficient (1) f | Partially (2) f | Sufficient (3) f |
|---|---|---|---|---|
| 1. The designed educational game corresponds to the learning outcomes mentioned. | 1 | 2 | 2 | 8 |
| 2. The purpose of the designed game is identified. | 0 | 0 | 4 | 9 |
| 3. The rules of the designed game are stated. | 1 | 0 | 2 | 10 |
| 4. Difficulty levels of the designed game correspond to the level of students. | 1 | 1 | 0 | 11 |
| 5. The designed educational game is well constructed in terms of time management. | 10 | 1 | 0 | 2 |
| 6. The designed educational game has an instructive feature. | 0 | 0 | 3 | 10 |
| 7. The designed educational game reinforces the new information related to the subject. | 0 | 1 | 3 | 9 |
| 8. The designed educational game can be used for evaluation. | 0 | 5 | 4 | 4 |
| 9. The designed educational game increases the social interaction among participants. | 0 | 2 | 3 | 8 |
| 10. The designed educational game creates a competitive relation among participants. | 0 | 4 | 2 | 7 |
| 11. The designed educational game creates an active learning environment for participants. | 0 | 1 | 2 | 10 |
| Total | 13 | 17 | 25 | 88 |

f: frequency

The initial item in evaluating the quality of the games was about the compatibility of the designed games with the mentioned learning outcomes. There was 1 design out of 13 (7.7%) that did not incorporate any learning outcome in this item, resulting in a score of zero. 2 designs (15.4%) received a score of 1 point, 2 designs (15.4%) received a score of 2 points, and 8 designs (61.5%) received a score of 3 points. The second item pertains to the purpose of the educational games that were designed. In this item, there were no designs that had a score 0 or 1. 4 designs (30.8%) that failed to state their purpose properly got 2 points. 9 designs (69.2%) received a score of 3 points for properly and sufficiently expressing their objective. The third item pertains to the rules of the designed games. One design (7.7%) that did not adhere to any rules received a score of zero. There were no designs that had a score of 1. Two designs (15.4%) got 2 points, while 10 designs (76.9%) received 3 points. The fourth item pertains to the difficulty level of the designed educational game and its compatibility with student's level. There was 1 design (7.7%) with 0 point, and one design (7.7%) with 1 point, but there were no designs with 2 points. 11 designs (84.6%) received a total of 3 points. The fifth item pertains to the time management of educational games designed. There were 10 designs (76.9%) with a 0-score. One of the designs (7.7%) received a score of 1 point. There were no designs with 2 points, but there were 2 designs (15.4%) that got 3 points each. The sixth item is about instructive feature of designed games. No design was assigned 0 and 1 points. 3 designs (23.1%) got 2 points and 10 designs (76.9%) got 3 points. The seventh item pertains to reinforcement of knowledge. No game was assigned 0 point. One game (7.7%) received a score of 1 point, 3 games (23.1%) received a score of 2 points, and 9 games (69.2%) received a score of 3 points. The eighth item is about using designed games for evaluation. No game was assigned 0 point. There were 5 games (38.4%) with 1 point, and 4 games (30.8%) with 2 points, 4 games (30.8%) with 3 points. The ninth item is about social interaction. No game was assigned 0 point. Two games (15.4%) received a score of 1 point, 3 games (23.1%) received a score of 2 points, and 8 games (61.5%) received a score of 3 points. The tenth item pertains to competition among participants of game. Four designs (30.8%) got 1 point, 2 designs (15.4%) got 2 points and 7 designs (53.8%) received 3 points. The eleventh item is about providing an active learning environment. No game got 0 point. There was 1 game (7.7%) with 1 point. There were 2 games (15.4%) with 2 points, 10 games (76.9%) with 3 points.

## PIE Evaluation

In this section, research findings related to the problem namely “While designing a game on environmental issues, how do the pre-service science teachers utilize the stages PIE?” were presented. The results are summarized in Table 6.

**Table 6** Findings Pertaining to the Compatibility of the Designed Games with the PIE Model

| Items | None (0) f | Insufficient (1) f | Partially (2) f | Sufficient (3) f |
|---|---|---|---|---|
| 1. What is the problem/issue/topic you would like to handle? | 0 | 0 | 7 | 6 |
| 2. What are your aims and objectives? | 0 | 0 | 1 | 12 |
| 3. What are the qualities and levels of the participants you will work with? | 1 | 0 | 1 | 11 |
| 4. What are the previous experiences, needs, interests and actions of the participants? | 2 | 0 | 1 | 10 |
| 5. What action and what kind of change are expected for each student? | 1 | 0 | 3 | 9 |
| 6. How can students be included in the planning process? | 0 | 1 | 2 | 10 |
| 7. Are there limitations/difficulties and/or resources? If yes, what are they? | 1 | 0 | 0 | 12 |
| 8. Which guidelines or activities do you think will have the greatest impact on the change in information, attitudes and behavior? | 0 | 1 | 5 | 7 |
| 9. After conducting pilot studies on the activities and materials to be implemented, what specific modifications have been identified? | 3 | 3 | 1 | 6 |
| 10. Is the plan, finance and employment prepared for the implementation suitable, sufficient and efficient? | 2 | 0 | 2 | 9 |
| 11. How would you know whether the method worked or not? | 0 | 0 | 2 | 11 |
| 12. What is the product or outcome of the program? | 0 | 0 | 2 | 11 |
| 13. Can you assess the metrics of achievement (attainment of objectives) such as the student's cognitive proficiency, disposition, and conduct, or alterations in the surroundings? | 0 | 2 | 4 | 7 |
| Total | 10 | 7 | 31 | 121 |

f: frequency

The games were evaluated based on the control list of PIE; the initial 8 items in the control list concern planning; items 9 and 10 concern implementation and items 11, 12 and 13 concern evaluation. The first item of the planning section of PIE pertains to the problems and topics to be handled in the designed games. There were no designs with a score value of 0 or 1. 7 designs (53.9%) acquired 2 points and 6 designs (46.1%) got 3 points. The second item concerns the aims and objectives. There were no designs with a score of 0 or 1. One design

(7.7%) got 2 points and 12 designs (92.3) got 3 points. The third item concerns the qualities and levels of participants in the games. There was only one design (7.7%) with a score of 0. There were no designs with 1 point; there was one design (7.7%) with 2 points and 11 designs (84.6%) with 3 points. The fourth item concerns the previous experiences, needs and actions. There were two designs (15.4%) with 0 point. There was no design with 1 point. One design (7.7%) got 2 points and 10 designs (76.9%) got 3 points. The fifth item concerns actions and changes made by each of the students. 1 design (7.7%) got 0 point. There was no design with 1 point. 3 designs (23.1) got 2 points and 9 designs (69.2%) got 3 points. The sixth item is about planning process. There was no design with 0 point. One design (7.7 %) got 1 point. 2 designs (15.4 %) got 2 points and 10 desings (76.9 %) got 3 points. The seventh item pertains to limitations and difficulties of designed games. There was only one design (7.7%) with 0 point. There were no designs with 1 and 2 points. 12 designs (92.3 %) got 3 points. The eighth item is about the greatest impact on change in attitude and behaviour. There was no design with 0 point. One design (7.7%) got 1 point. 5 designs (38.4%) got 2 points and 7 designs (53.8%) got 3 points. The nineth item is about changes after pilot implementation. 3 designs (23.1%) got 0 point. There were 3 designs (23.1%) with 1 point, 1 design (7.7%) with 2 points, 6 designs (46.1%) with 3 points. The tenth item pertains to the plan and finance for the implementation. 2 games (15.4%) got 0 point. There was no game with 1 point. 2 games (15.4%) got 2 points and 9 games (69.2%) got 3 points. The eleventh item evaluates whether the method works. There were no with design 0 and 1 points. 2 designs (15.4%) got 2 points and 11 designs (84.6%) got 3 points. The twelfth item is about the product or outcome of the program. There was no design with 0 and 1 points. There were 2 games (15.4%) with 2 points and 11 games (84.6%) with 3 points. The thirteenth item is about assessment of the metrics of achievement. There was no design with 0 point. 2 designs (15.4%) got 1 point. 4 designs (30.8%) got 2 points and 7 desings (53.8%) got 3 points.

## Environmental Literacy

The games designed by the pre-service teachers were examined in the framework of environmental literacy components through the research question “Which components of environmental literacy are employed by the pre-service science teachers while designing games based on the PIE model?’’. Which component was included mostly in the games by the pre-service teachers is presented in Table 7.

**Table 7** Environmental Literacy Components

| Environmental literacy components | Number of games included the component | Number of games did not include the component |
|---|---|---|
| Environmental awareness | 13 | 0 |
| Environmental knowledge | 12 | 1 |
| Environmental attitude | 11 | 2 |
| Environmental skill | 1 | 12 |
| Environmental behavior | 8 | 5 |

When data in Table 7 were examined, the graphs demonstrated that all of the pre-service teachers employed environmental awareness among the components of environmental literacy, while utilizing environmental skills (7.7%) the least. After the environmental awareness, there came environmental knowledge (92.3%), environmental attitude (84.6%) and environmental behavior (61.5%) respectively.

## Discussion

The study evaluated the educational games created by the pre-service science teachers based on their general features, quality, compatibility with the PIE model, and inclusion of environmental literacy components. Initially, when examining the attributes of games, it has been observed that games were typically intended for group play. Due to the overcrowded nature of educational system in Türkiye, students face increased challenges in participating in individual or extracurricular games, primarily due to constraints in time management and adherence to game rules. One possible explanation for why the pre-service teachers favor group games could be due to this factor. Regarding the type of games that were created, it has been discovered that the majority of them were card games. When Cop and Kablan (2018) analyzed the studies on the educational games in Türkiye, it was seen that the mostly used games in the studies were computer-based games. Nevertheless, not all teachers possess the expertise and resources to create a computer game. One possible explanation for the prevalence of card game design in this study is that it does not necessitate specialized knowledge or skills, such as "technology". Most of the games designed by the pre-service teachers, the components, such as objectives, feedback, rewards, and rules, were appropriately defined. Furthermore, these games were designed to be engaging and capable of being played multiple times. In the study by Önen et al. (2012), most of the games designed by the pre-service science teachers had rules defined clearly and there was a special care for games' being enjoyable. At this point, it can be said that the pre-service teachers had positive skills and knowledge related to the usage of games in science education.

Upon reviewing the literature on educational games, numerous studies have discovered that a key advantage of incorporating games in education is the active engagement of students in the classroom (Çakır, 2022; Dolunay & Karamustafaoğlu, 2021; Kılıç & Karamustafaoğlu, 2020; Yıldız et al. 2016). Upon examining the attributes of the games created by the participants in this study, it was discovered that the games they designed excelled in terms of "facilitating an engaging educational setting". As well as games, PIE design is also instrumental in engaging students. Keleş et al. (2016) characterized this model as a means of promoting student engagement, enhancing learning outcomes, analyzing performance-based issues, and utilizing the steps of planning, implementation and evaluation. Moreover, the qualities of the games designed by the pre-service teachers have been evaluated as sufficient in terms of these criteria, as the learning outcomes, class level, objectives and rules were specified in most of the games. Nevertheless, due to the absence of a predetermined timetable for game sessions, the designed games have been deemed insufficient in terms of time allocation and organization. Another crucial aspect of the games developed in this study was their ability to augment social interaction and foster competition among participants. Lee et al. (2011) found interactive educational games involving social interaction and competition to be more effective in terms of learning processes compared to individual offline ones. The games designed were also found to be sufficient for instruction and reinforcement, but insufficient for evaluation purposes. Earlier studies on educational games have also shown that educational game is used more as a teaching technique (Fırat, 2011; Karamastafaoğlu & Baran, 2020; Kılıç & Karamustafaoğlu, 2020), while it is used less for evaluation purposes (Yıldırım & Can, 2017). Nevertheless, games have a significant potential to incorporate the utilization of knowledge and abilities within the educational setting. In an environment characterized by motivated students who engage in enjoyable activities to assess their knowledge and skills, teachers can conduct evaluations, and students' self-evaluations can also be valuable.

Upon evaluating the designed games in relation to PIE components, it was found that they were sufficient in terms of many components within the framework. The implementation of a question-and-answer paper, with clearly defined PIE components, as a means of collecting data from the pre-service teachers has resulted in a high level of satisfaction with the participants' responses. Given that PIE's framework is a specialized model tailored for pre-service teachers interested in environmental game design, it is anticipated that these pre-service teachers will achieve success in this field.

When the games designed by the pre-service teachers were examined, it was discovered that the most prevalent environmental literacy component was awareness, followed by knowledge and attitude while the skill components have been identified in a single design. There have been two different interpretations of this result. One possibility is that the pre-service teachers may have incorporated similar components drawing from their own environmental literacy. When studies on this topic were examined, it was discovered that the pre-service teachers generally exhibited a high level of environmental attitude (Özgürler, 2014; Şahin et al., 2016) but a low level of knowledge and behavior (Altınöz, 2010; Bilim, 2012; Özgürler, 2014; Şahin et al., 2016; Teksöz et al., 2010). The high level of attitude observed in the studies and the pre-service teachers' involving environmental attitude in higher rates in their designed games aligns with our interpretation. Moreover, environmental knowledge and awareness were utilized in this study with high frequency. It was perceived as a positive reflection of the game designs because they were developed in an environmentally focused class in a term when SSE was taught.

In addition to these findings, the utilization of the environmental skill component in game designs was observed to be the least frequent. An analysis of learning outcomes in the 2018 science curriculum (MoNE, 2018) in terms of environmental literacy revealed that the learning outcomes related to knowledge and skill were mostly included in game designs, and the learning outcomes related to awareness, attitude and behavior were not included enough. (Fidan Yazgan, 2023). As the pre-service teachers adhered to the 2018 curriculum when creating their designs, it was observed that there was consistency in the learning outcomes related to knowledge, however there was a difference in the skill component. In this reference, it can be said that environmental skill component included in the environmental learning outcomes of the science education curriculum was not much adhered too much. This may be associated with the complexity of environmental problems and the utilization of multidiscipline for their solution. Here in, it can be asserted that the pre-service teachers created their designs based on their professional qualifications, individual differences, and their own environmental knowledge and perspectives, rather than the achievements they discussed.

## Conclusions and Suggestions

Upon considering all the findings, it was seen that pre-service teachers did not encounter too many problems with the game components in general while designing games. According to the findings of the study, it was noticed that the pre-service teachers did not include environmental literacy components sufficiently in their game designs, especially skills and behaviour component. Based on this, it can be suggested to emphasize environmental literacy components during the course. For this purpose, the methods or contents of the course may be modified. For this study, it may be a limitation that only the reports submitted by the pre-service teachers were used as data collection tools. The use of observation forms or conducting interviews with participants may provide more detailed information for further studies. Additionally, providing pre-service teachers an opportunity to apply the designed games in the schools with students may affect their professional development positively. Because of the attractive nature of fames for students, introducing elective courses for pre-service teachers at the faculties of education, focusing on how to design a game and its integration into their classes can enhance their professional development.

## Compliance with Ethical Standards

*Disclosure of potential conflicts of interest*

Authors declare that they have no competing interest

*Funding*

This research did not receive any funding from agencies

*CRediT author statement*

The fourth author implemented the study. All authors took part in writing the article. First author also made editing and reviews as corresponding author.

*Research involving Human Participants and/or Animals*

This study was carried out taking ethical rules into account.

## Fen Eğitiminde Planlama, Uygulama, Değerlendirme Modeline Uygun Oyun Tasarımlarının İncelenmesi

**Özet:**

Bu çalışmanın amacı, fen bilgisi öğretmen adayları tarafından PIE (Planla, Uygula, Değerlendir) modeline göre tasarlanan oyunları; genel özellikleri, niteliği, PIE modeline uygunluğu ve çevre okuryazarlığı bileşenleri açısından incelemektir. Bu amaçla 3. Sınıfta öğrenim gören ve Fen, Sürdürülebilirlik ve Çevre (FSÇ) seçmeli dersini alan fen bilimleri öğretmen adayları araştırmaya katılmıştır. Öğretmen adaylarından Millî Eğitim Bakanlığı 2018 Fen Bilimleri Öğretim Programındaki çevre konulu kazanımlar ele alınarak PIE modeline göre oyun tasarımı yapmaları istenmiştir. Tasarımlar üç araştırmacı tarafından birbirinden bağımsız olacak şekilde değerlendirilmiş ve Miles ve Huberman'in (1994) sunduğu uyuşum formülü ile güvenirliğine bakılmıştır. Çalışma sonucunda oyunların çoğunlukla kart oyunu türü olduğu ve özgün tasarım oldukları görülmüştür. Oyunların nitelik bakımında yeterli ve PIE modelinin çerçevesine uygun tasarımlar olduğu sonucuna ulaşılmıştır. Öğretmen adaylarının oyunlarında çevre okuryazarlık bileşenlerinden en çok farkındalık ardından da bilgi bileşenlerine yer verdikleri tespit edilmiştir.

*Anahtar kelimeler:* Eğitsel oyunlar, fen eğitimi, PIE modeli, çevre okuryazarlığı.

## References

Akcanca, N., & Sömen, T. (2018). Prospective teachers' educational game designing and applications. *Electronic Turkish Studies*, *13*(27), 49-71. http://dx.doi.org/10.7827/TurkishStudies.14506

Akgül, G., & Kılıç, M. (2020). Preservice science teachers' views on educational digital games and KODU implementations. *Fen Bilimleri Öğretimi Dergisi*, *8*(2),101-120. https://dergipark.org.tr/tr/pub/fbod/issue/71992/1158003

Alınmaz, N. (2023). *Investigation of pre-school teachers' competencies on environmental literacy* (Publication No. 774963) [Master's thesis, Kastamonu University]. Council of Higher Education Thesis Center.

Altınöz, N. (2010). *Environmental literacy levels of prospective science teachers* (Publication No. 274704) [Master's thesis, Sakarya University]. Council of Higher Education Thesis Center.

Andrews, D. H., & Goodson, L. A. (1980). A comparative analysis of models of instructional design. *Journal of Instructional Development*, *3*(4), 2-16. https://link.springer.com/article/10.1007/BF02904348

Aymen Peker, E. (2018). *Evaluation of teaching the unit of titled 'Let's travel and learn about the living world' for 5th grades with classic educational games and educational games supported with technology* (Publication No. 492301) [Doctoral dissertation, Ondokuz Mayıs University]. Council of Higher Education Thesis Center.

Beker Baş, M., & Karamustafaoğlu, O. (2020). An educational game on central and peripheral nervous system. *Gazi Journal of Educational Science, 6*(1), 80-92.

https://dergipark.org.tr/tr/pub/gebd/issue/53429/643014

Benzer, E. (2010). *The effect of environmental education lesson prepared with project-based learning approach on pre-service science teachers' environmental literacy* (Publication No. 279892) [Doctoral dissertation, Marmara University]. Council of Higher Education Thesis Center.

Bilim, İ. (2012). *Sustainable environment perspective faculty of education students' levels of environmental literacy* (Publication No. 330235) [Master's thesis, Afyon Kocatepe University]. Council of Higher Education Thesis Center.

Cop, M. R., & Kablan, Z. (2018). The analysis of studies on educational games in Turkey. *Kocaeli University Journal of Education*, *1*(1), 52-71. https://dergipark.org.tr/tr/pub/kuje/issue/37623/422759

Çakır, A. (2022). *The effect of educational games developed for the third grade science course five sense unit on the academic success of students* (Publication No. 770891) [Master's thesis, Amasya University]. Council of Higher Education Thesis Center.

Dolunay, A., & Karamustafaoğlu, O. (2021). Opinions of science teachers about educational games: "The fastest sound" game". *International Journal of Turkish Educational Sciences*, *2021*(16), 48-69.

https://dergipark.org.tr/tr/pub/goputeb/issue/62038/872860

Durdut, A. (2016). *Improving an educational computer game in mathematics for vocational school students and the effects of success* (Publication No. 456152) [Master's thesis, Balıkesir University]. Council of Higher Education Thesis Center.

Fang, W. T., Hassan, A. A., & LePage, B. A. (2022). *The living environmental education: Sound science toward a cleaner, safer, and healthier future*. Springer.

Fırat, S. (2011*). The effects of mathematics teaching performed through educational computer games on conceptual learning* (Publication No. 301095) [Doctoral dissertation, Adıyaman University]. Council of Higher Education Thesis Center.

Fidan Yazgan, P. (2023). *Comparison of the objectives in the national science curriculum and optional environmental education programs in the context of environmental literacy* (Publication No. 783659) [Master's thesis, Marmara University]. Council of Higher Education Thesis Center.

Freitas, S., & Oliver, M. (2006). How can exploratory learning with games and simulations within the curriculum be most effectively evaluated?. *Computers & Education*, *46*(3), 249-264. https://doi.org/10.1016/j.compedu.2005.11.007

Gee, J. P. (2003). What video games have to teach us about learning and literacy?. *Computers in Entertainment*, *1*(1), 1-4. http://dx.doi.org/10.1145/950566.950595

Herodotou, C. (2018). Mobile games and science learning: A comparative study of 4 and 5 years old playing the game Angry Birds. *British Journal of Educational Technology*, *49*(1), 6-16. https://doi.org/10.1111/bjet.12546

Jacobson, S. K., McDuff, M. D., & Monroe, M. C. (2006). *Conservation education and outreach techniques (Techniques in ecology and conservation).* Oxford University Press.

Karamustafaoğlu, O., & Baran, S. (2020). Teachers' views about science teaching with 'Force capturing' the educational game. *Fen Bilimleri Öğretimi Dergisi*, *8*(1), 76-91. https://dergipark.org.tr/tr/pub/fbod/issue/71986/1158007

Karamustafaoğlu, O., & Kılıç, M. F. (2020). Investigation of national scientific studies about educational games. *Atatürk University Journal of Kazım Karabekir Education Faculty*, (40), 1-25. https://dergipark.org.tr/tr/pub/ataunikkefd/issue/55385/730393

Kefeli, N., Taş, E. & Yalçın, M. (2018). The effect of teaching science with word games on the attitudes of students towards environment. *International e-Journal of Educational Studies*, *2*(3), 44-52. https://doi.org/10.31458/iejes.388619

Keleş, E., Fiş Erümit, S., Özkale, A., & Aksoy, N. (2016). A roadmap for instructional designers: A comparison of instructional design models. *Ankara University Journal of Faculty of Educational Sciences*, *49*(1), 105-139. https://doi.org/10.1501/Egifak_0000001377

Kılıç, M., & Karamustafaoğlu, O. (2020). Teachers' view on blood catching educational game. *Journal of Science Mathematics Entrepreneurship and Technology Education*, *3*(3),144-158. https://dergipark.org.tr/tr/pub/fmgted/issue/60204/838898

Kırmızıyüz, E., Ercan, D., & Uz Bilgin, Ç. (2021). Designing, developing and implementing an educational mobile game on electrical energy conversion and recycling. *Yıldız Journal of Educational Research*, *6*(1), 48-60. https://doi.org/10.51280/yjer.2021.013

Kirriemuir, J., & McFarlane, A. (2003, January). Use of computer and video games in the classroom [Conference Paper] Proceedings of DiGRA 2003 Conference: Level Up, Utrecth, Holland. https://www.researchgate.net/publication/221217309_Use_of_Computer_and_Video_Games_in_the_Classroom

Kocadere, S. A., Özhan, Ş. Ç., Bayrak, F. & Kibar P.N. (2019). Story of Hercules: An educational game development model proposal. *Educational Technology Theory and Practice*, *9*(1), 230-250. https://doi.org/10.17943/etku.451990

Korkmaz, S. (2018). *Development and effectiveness of educational games in science teaching as supplementary material on students' achievement and attitude towards science* (Publication No. 502718) [Master's thesis, Bartın University]. Council of Higher Education Thesis Center.

Külegel, S. (2020). *A research to improve the 21st century skills of gifted and talented students of environmental education-based science, technology, engineering, mathematics activities* (Publication No. 640643) [Master's thesis, Yıldız Teknik University]. Council of Higher Education Thesis Center.

Lee, K. M., Jeong, E. J., Park, N., & Ryu, S. (2011). Effects of interactivity in educational games: A mediating role of social presence on learning outcomes. *International Journal of Human–Computer Interaction*, *27*(7), 620-633. https://doi.org/10.1080/10447318.2011.555302

Miles, M. B., & Huberman, A. M. (1994). *Qualitative data analysis: An expanded sourcebook* (2nd ed.). Sage.

Ministry of Education (MoNE) (2018). İlköğretim kurumları (ilkokullar ve ortaokullar) fen bilimleri dersi (3,4,5,6,7 ve 8. sınıflar) öğretim programı. https://mufredat.meb.gov.tr/Dosyalar/201812312311937-FEN%20B%C4%B0L%C4%B0MLER%C4%B0%20%C3%96%C4%9ERET%C4%B0M%20PROGRAMI2018.pdf

Mutlu Karanfil, B. (2022). *Investigation of the effect of portfolio-based environmental education with family participation on environmental awareness of 5th grade students* (Publication No. 708253) [Master's thesis, Trakya University]. Council of Higher Education Thesis Center.

Önen, F., Demir, S., & Şahin, F. (2012). Science teacher candidates' views about games and evaluating prepared games by them. *Ahi Evran University Journal of Kırşehir Education Faculty*, *13*(3), 299-318. https://dergipark.org.tr/tr/pub/kefad/issue/59486/854950

Özen, G., Timurkaan Akçınar, S., Güllü, M., Meriç, F., Uğraş, S., & Çelik Çoban, D., (2019). Spor liseleri 12. sınıf eğitsel oyunlar ders kitabı. Milli Eğitim Bakanlığı.

Özgürler, S. (2014). Investigation of environmental literacy level and ecological footsprint of teacher candidates (Publication No. 380580) [Master's thesis, Amasya University]. Council of Higher Education Thesis Center.

Özkan, G., Akça, E., & Umdu Topsakal, Ü., (2020). The effect of game-based teaching on students' attitudes towards the science course and views. *The Journal of International Social Research*, *13*(70), 794-800. http://dx.doi.org/10.17719/jisr.2020.4135

Serdaroğlu, C., & Güneş, M. H. (2019). Effects on academic achievement and attitude of game based learning in teaching of reproduction, growth and development unit in 6th grade plant and animal. *Necatibey Faculty of Education Electronic Journal of Science and Mathematics Education, 13*(2), 1015-1041. https://doi.org/10.17522/balikesirnef.570616

Şahin, S., Ünlü, E., & Ünlü, S. (2016). Investigation of teacher candidates' environmental literacy awareness level. *Education Sciences*, *11*(2), 82-95. https://dergipark.org.tr/en/pub/nwsaedu/issue/19844/212575

Seçkin Kapucu, M., & Çağlak, S. (2018). Pre-service science teachers' skills and experience on designing educational games: A case study. *Adıyaman University Journal of Social Sciences*, (29), 536-573. https://doi.org/10.14520/adyusbd.364051

Şentürk, C. (2020). The effects of game-based science learning experiences on academic achievement, retention, attitudes and learning process. *The Journal of National Education*, *49*(227), 159-194. https://dergipark.org.tr/tr/pub/milliegitim/issue/56322/616595

Şimşek, A. (2018). *Sosyal bilimlerde araştırma yöntemleri*. Anadolu Üniversitesi.

Teksöz, G., Şahin, E., & Ertepınar, H. (2010), Environmental literacy, pre-service teachers and a sustainable future. *Hacettepe University Journal of Education*, *39*, 307-320. https://dergipark.org.tr/tr/pub/hunefd/issue/7799/102184

Yıldırım, A., & Şimşek H. (2021). *Sosyal bilimlerde nitel araştırma yöntemleri*. Seçkin.

Yıldırım, M., & Can, S. (2017). "Deal or no deal": Should science lesson be taught with instructional games?. *Atatürk University Journal of Kazım Karabekir Education Faculty*, *35*,14-30. https://dergipark.org.tr/tr/pub/ataunikkefd/issue/33367/336640

Yıldız, E., Şimşek, Ü., & Araz, H. (2016). The effect of the educational game method on academic achievement and motivation towards science learning in teaching of circulatory system. *Mustafa Kemal University Journal of Social Sciences Institute*, *13*(36) 20-32. https://dergipark.org.tr/tr/pub/mkusbed/issue/44301/547052

Necatibey Faculty of Education Electronic Journal of Science and Mathematics Education
Vol. 18, Issue 2, December 2024, pp. 248-270. ISSN: 1307-6086

Research Article

# The Effect of Experiences of Comparing International Mathematics Curricula on Prospective Mathematics Teachers' Views on Primary Mathematics Curriculum

**Hakan ULUM**[1]

[1] Necmettin Erbakan University, Ereğli Faculty of Education, Türkiye, hakanulum@gmail.com, https://orcid.org/0000-0002-1398-6935

Received : 01.04.2024 Accepted : 24.07.2024

Doi: https://doi.org/10.17522/balikesirnef.1462714

---

*Abstract* – This research was conducted among prospective mathematics teachers during the 2023-2024 academic year at a university in the Central Anatolia Region of Türkiye. The study's primary objective was to investigate the impact of comparative experiences with international mathematics curricula on the perceptions and attitudes of prospective mathematics teachers towards the elementary mathematics curriculum in Türkiye. Specifically, the study aimed to explore how the experiences gained by prospective mathematics teachers through the examination of international mathematics curricula reflected their perceptions and attitudes towards their country's elementary mathematics curriculum. In the study, a one-group pretest-posttest experimental design was used and supported by interviews. The sample consisted of 10 prospective mathematics teachers enrolled in an elective course called "Comparison of Mathematics Teaching Programs." The findings indicate that the experience of analysing international mathematics curricula positively influenced the participants' perceptions and attitudes towards teaching programmes. Prospective teachers suggested improvements such as developing a curriculum closer to international standards, increasing student-centeredness, and diversifying assessment and evaluation methods. The results provide significant insights for developing elementary mathematics curricula and underscore the importance of integrating global perspectives into teacher education programs.

*Keywords*: Mathematics curriculum, preservice teachers, experimental design.

------------

Corresponding author: Hakan ULUM, hakanulum@gmail.com

## Introduction

Mathematics education is a critical aspect of students' academic development, and the curriculum plays a vital role in shaping students' views and understandings of the subject (Wößmann, 2003). Experiences of comparing international mathematics curricula can profoundly impact individuals' views and perceptions of the mathematics curriculum. Comparing different international mathematics curricula can provide opportunities for teachers and educators to gain insights into other approaches, methods and content covered in mathematics education (Atweh & Clarkson, 2002). This can broaden their understanding of effective teaching strategies and help them reflect on the strengths and weaknesses of their curricula.

Furthermore, comparing international mathematics curricula can highlight areas where curricula may be lacking or improved (Tatto & Senk, 2011). By examining the similarities and differences between different curricula, educators can identify gaps or deficiencies in their curricula and try to address them (Atweh & Clarkson, 2002). Moreover, comparing international mathematics curricula can foster collaboration among mathematics educators. Teachers and educators can participate in discussions and exchanges with global colleagues to share experiences, exchange ideas and learn from each other's expertise. This collaboration can develop a more comprehensive and well-rounded mathematics curriculum incorporating best practices from different countries (Cai et al., 2017; Kaiser et al., 2002; Zhou et al., 2020). Overall, experiences comparing international mathematics curricula can have a transformative effect on individuals' views, leading to a deeper understanding of effective teaching methods and an awareness of areas for improvement in their curricula.

Consequently, experiences comparing international mathematics curricula have the potential to foster a collaborative and innovative ecosystem that crosses geographical boundaries and ultimately shapes the future of mathematics education. Furthermore, studying international mathematics curricula can challenge the notion of a global curriculum and instead encourage an international approach to curriculum reform (Atweh & Clarkson, 2002). As the global environment continues to evolve, the field of mathematics education faces new challenges and opportunities. The comparison of international mathematics curricula has emerged as an area of focus of interest for educators seeking to enrich teaching methods and enhance students' learning experiences.

## General Overview of Countries' Mathematics Curricula

Mathematics curricula in different countries vary in structure, content and emphasis. Countries like the United States and China have curricula focusing more on computational skills and algorithms, while Finland and Singapore may prioritise problem-solving and critical thinking (Hjalmarson, 2008). Within these curricula, there are also different views on the role of play in mathematics education for young children. In Finland, for example, play is essential to early childhood mathematics education, enabling children to understand mathematical concepts through hands-on exploration (Peter-Koop & Scherer, 2012). In contrast, countries like Japan may have a more formal and structured approach to teaching mathematics at an early age.

A comparison of international mathematics curricula reveals various approaches and emphases. For example, while some countries, such as Japan and South Korea, emphasise rote learning and memorisation in mathematics education, others, like the United States and Finland, prioritise conceptual understanding and applying mathematical principles in the real world (Atweh & Clarkson, 2002). The use of technology in mathematics education also varies across countries. Some countries, such as Indonesia and China, emphasise using technology to understand and simulate abstract ideas in mathematics.

In addition to differences in content and teaching methodologies, the rigour and depth of mathematics curricula also vary from country to country. For example, countries like Singapore and South Korea introduce advanced mathematical concepts earlier, while the United States and Finland may adopt a more gradual and progressive approach to introducing complex topics (KERIS, 2020; Ministry of Education, Singapore, 2020; Nguyen et al., 2019; US Department of Education, 2010). Regarding teachers' readiness and competence in teaching mathematics, there are also different levels of training and preparation across countries. Nations like Japan and Germany have well-trained, confident educators who prioritise mathematics teaching (KMK, 2020; Ministry of Education, Culture, Sports, Science and Technology, Japan, 2020). In contrast, others may need more prepared educators and more confidence in effectively teaching the subject (Clark-Wilson et al., 2020).

Moreover, the role of technology in mathematics education varies across countries. For instance, countries like the United States and Singapore use technology and digital tools as integral components of their mathematics curricula. At the same time, Japan and Germany may maintain a more traditional approach based primarily on paper-and-pencil methods (Atteh et al., 2020). Overall, the mathematics curricula of different countries exhibit a variety of techniques, ranging from play-based exploration to formal instruction and from an emphasis on rote memorisation to conceptual understanding and real-world application. It is important to note that each country's mathematics curriculum is shaped by cultural, educational, and societal factors specific to that country (Atteh et al., 2020; Zhang et al., 2018).

Recognising and understanding these differences is crucial to comprehensively understanding how mathematics is taught and learned worldwide. By acknowledging the diversity of approaches, educators and policymakers can draw inspiration from best practices in different countries to improve mathematics education in their education systems (Clark-Wilson et al., 2020; Zhang et al., 2018). Moreover, by building bridges between different approaches and experiences, teachers can promote equity and educational effectiveness in mathematics education. Mathematics education is complex and dynamic; technology integration is essential for successful teaching.

## Aim of the Study

This study examines the effects of comparing international mathematics curricula on pre-service mathematics teachers' views of the primary mathematics curriculum and how these experiences change and shape their perceptions.

The problems of the study are as follows;

1. Do experiences comparing international mathematics curricula significantly change preservice mathematics teachers' views of the elementary mathematics curriculum?

2. How do experiences comparing international mathematics curricula shape preservice mathematics teachers' views of the primary mathematics curriculum?

This study aims to understand the effects of the experiences of comparing international mathematics curricula on preservice mathematics teachers' views of the primary mathematics curriculum. In other words, the study's primary purpose was to investigate how preservice mathematics teachers' experiences examining international mathematics syllabi reflected their perceptions of their home country's primary mathematics curriculum and how these experiences affected their attitudes towards the curriculum. For this purpose, an in-depth analysis of the participants' experiences evaluating international mathematics syllabi and the effects of these experiences on local mathematics curricula can be conducted.

## Method

### Research Design

This study used a one-group pretest-posttest experimental design with a combination of quantitative research methods to determine how the experiences of comparing mathematics curricula changed and shaped pre-service mathematics teachers' perceptions of the elementary mathematics curriculum. Interviews supported quantitative research findings. This design aims to provide a more detailed and comprehensive understanding of a phenomenon (Mills & Gay, 2016). With the one-group pretest-posttest experimental design, the statistical significance between the pre-test and post-test scores of the participant group will be determined with the opportunity to evaluate the effect of the application in a single group (Büyüköztürk et al., 2011). Although the one-group pretest-posttest experimental design is generally accepted as a weak design in research, it is considered a preference that should be used when a new training module is developed and implemented. In this study, explaining the cause-and-effect relationship between dependent and independent variables was preferred when a new training module was implemented and evaluated (Creswell, 2018).

In addition to quantitative data, interviews were conducted to obtain in-depth information. Interviews were collected to understand pre-service mathematics teachers' perceptions of the primary mathematics curriculum based on their experiences comparing mathematics curricula (Patton, 2014). In this way, a more holistic understanding of pre-service mathematics teachers' perceptions and experiences of the primary mathematics curriculum was attempted based on their experiences of comparing curricula.

### Participants

The study's population is prospective mathematics teachers studying in Türkiye in the 2023-2024 academic year. Within this population, the sample consisted of 10 preservice mathematics teachers who took the elective course "Comparison of Mathematics Curricula" at a university in the Central Anatolia Region of Türkiye in the 2023-2024 academic year. Therefore, the study will focus on the experiences and perceptions of this sample of preservice mathematics teachers. The selection of this sample allows for an in-depth examination of a specific geographical and institutional context.

### Material

The scale form described below and a semi-structured interview form prepared by the researcher will be used in the study.

### *Mathematics programme evaluation scale according to the expert-oriented programme evaluation approach (MPAS)*

The Mathematics Curriculum Evaluation Scale, designed based on the expert-oriented curriculum evaluation model developed by Yıldız and Gürgen (2021), will be used in the study. In the process of creating the scale, a draft consisting of 65 items was prepared as a result of the literature review, and this draft was submitted to the evaluation of 4 experts specialised in the field of curriculum and measurement and assessment, as well as three classroom teachers. In line with the expert and teacher opinions, the draft was reduced to 40 items and transformed into a scale form using a 5-point Likert-type scale. The scale consists of 40 items, including the programme's objectives, topics, teaching-learning process and measurement-evaluation dimensions. Participants indicated their opinions by marking one of the options "Strongly Agree (5), Agree (4), Partially Agree (3), Disagree (2) and Strongly Disagree (1)" for each item. The scale's reliability was calculated using the Cronbach Alpha (.95) internal consistency coefficient, and reliability values were obtained separately for different sections. Cronbach Alpha internal consistency coefficient of the scale for this study was calculated as .89. At the beginning of the scale, general information about the purpose of the research and instructions for filling out the questionnaire were given. At the beginning of each section, specific explanatory statements were included. This structure was added to ensure the participants correctly understood and completed the scale.

### *Semi-structured interview form*

The other data collection tools used in the study were the semi-structured interview form and the written opinions of the preservice teachers. The researchers prepared a semi-structured interview form that consisted of five questions. After the interviews with the preservice teachers, the participants submitted a written form to the researchers reflecting their general opinions about comparing mathematics curricula.

## Data Collection

The relevant scales used in the study were collected from university students using a written form. In the interviews conducted using a semi-structured interview form, voice recordings were taken with the participant's consent. Afterwards, the voice recordings were converted into written text to prepare them for coding. The collected data were transferred to the computer environment and stored.

## Data Analyses

Descriptive statistical analysis, t-tests and effect size calculations were used to analyse the quantitative data. Before statistical analysis, the data were checked for normality and descriptive statistics. The analysis of the interview recordings aimed to reveal the experience of comparing international mathematics curricula realistically and holistically based on pre-service mathematics teachers' perceptions. The participants were regarded as the data source, and the data obtained were organised, categorised and coded (Patton, 2014).

## Validity and Reliability

Validity and reliability measures were taken within the scope of the research. The reliability coefficients of the data obtained from the scales were calculated. During the research process, the researcher, who is the instructor of the elective course, adopted the principle of moving away from his prejudices and preconceived ideas about international mathematics curricula. A neutral approach was adopted to prevent the prior knowledge obtained from the literature review from influencing the research.

Four independent coders participated in the data analysis and reached a consensus regarding the coding process. All statements were reviewed for accuracy. that would direct the participants were avoided during the interviews. The data were presented to the reader through direct quotations without adding comments. Relevant comments and harmful data were also presented. Therefore, reliability and validity measures were taken during the research process to ensure the qualitative data were analysed accurately and reliably.

## Experimental Implementation Process

Comparison of Mathematics Curricula" elective course was conducted in a weekly 2-hour schedule for 14 weeks. In the first week of the course, an introduction was made to compare mathematics curricula, and information was given about the history and basic principles of different mathematics curricula. The (practical) applications began from the second week onwards, and examples from different mathematics curricula (Singapore, Finland, Japan, and the United States) were presented to the students. We examined these curricula each week by comparing them with the Turkish Mathematics Education Programme. We critically engaged with each country's programme through presentations, discussions, and Q&A activities. In Weeks 2 and 3, environments were created for students to do free work and exploration, and the characteristics of different mathematics curricula were focused on.

In weeks 4-13, the Ministry of National Education Primary 1-8. Grade Mathematics Curriculum (2018) was compared with the curricula of different countries, such as Singapore, Finland, Japan, and the United States. These countries were selected based on their success in international mathematics exams. The preservice teachers in the classroom presented the comparative studies and explained their similarities and differences with our mathematics curriculum. As a result of these comparisons, discussions were held on how to enrich our curriculum, and the strengths and weaknesses of different mathematics curricula were discussed. In the last week of the course, a general evaluation was made, and the knowledge and experiences gained by the students were reviewed.

Before the course, data were collected using the scales determined within the study's scope and a semi-structured interview form. The scales were administered on the 14th.

## Findings

### Findings Related to Quantitative Measurements

The central tendency measures, skewness, and kurtosis coefficients of the data obtained within the research scope were examined to check whether the data showed a normal distribution. It was accepted that the data showed a normal distribution. In addition, this acceptance was validated by testing with the Shapiro-Wilk test. The results obtained are presented in Table 1.

**Table 1** Shapiro-Wilk Test Results of MPDS Scale Data

| | Statistics | SD | p |
|---|---|---|---|
| MPDS scale | .989 | 9 | .996 |

According to Table 1, the data obtained from the scales show a normal distribution ($p>0.05$). In this context, parametric tests were applied to the data. The study analysed whether there was a significant difference between the mean scores of both scales of the participants using the related sample t-test technique. The results are given in Table 2.

**Table 2** Paired Sample T-test Results of Pre and Post-test Mean Scores of Preservice Teachers' MPDS Scale

| | | n | X | s | SD | t | p* |
|---|---|---|---|---|---|---|---|
| Attitudes towards the outcomes of primary mathematics curriculum | Pre-test | 10 | 23.90 | 4.43 | 9 | -3.713 | .005 |
| | Post-test | 10 | 26.10 | 3.07 | | | |
| Attitudes towards the content of primary mathematics curriculum | Pre-test | 10 | 24.70 | 3.83 | 9 | -2.496 | .034 |
| | Post-test | 10 | 26.20 | 2.78 | | | |
| Attitudes towards the learning-teaching processes of the primary mathematics curriculum | Pre-test | 10 | 24.20 | 2.65 | 9 | -4.392 | .002 |
| | Post-test | 10 | 25.70 | 1.70 | | | |
| Attitudes towards the measurement and evaluation dimension of the primary mathematics curriculum | Pre-test | 10 | 23.30 | 2.40 | 9 | -2.946 | .016 |
| | Post-test | 10 | 25.10 | 2.55 | | | |
| Attitudes towards primary mathematics curriculum | Pre-test | 10 | 96.10 | 7.21 | 9 | -7.826 | .001 |
| | Post-test | 10 | 103.10 | 5.83 | | | |

*p<0.05

In a class of 10 students in which the effect of the "Comparison of Mathematics Curricula" course, which is based on the comparison of the mathematics curricula of the countries, on the attitudes of prospective mathematics teachers towards the primary mathematics curriculum was investigated, as a result of the t-test for related samples conducted to determine whether there was a difference between the mean scores of the attitudes towards the primary mathematics curriculum applied at the beginning and end of the course;

It was found that the difference between the mean score of the scale applied at the beginning of the course (pretest = 23.90). The mean score of the scale applied at the end of the course (post-test = 26.10) for the sub-dimension of attitudes towards the objectives of the elementary mathematics curriculum was significant [$t(9) = -3.713$, $p < 0.05$].

For the sub-dimension of attitudes towards the content of the elementary mathematics curriculum, the difference between the mean score of the scale applied at the beginning of the course (pretest = 24.70) and the mean score of the scale applied at the end of the course (post-test = 26.20) was found to be significant [$t(9) = -2.496$, $p < 0.05$].

It was determined that the difference between the mean score of the scale applied at the beginning of the course (pretest = 24.20) and the mean score of the scale applied at the end of the course (post-test = 25.70) for the attitudes towards the learning-teaching processes sub-dimension of the primary mathematics curriculum was significant [$t(9) = -4.392$, $p < 0.05$].

The difference between the mean scale score applied at the beginning of the course (pretest = 23.30) and the mean score of the scale used at the end of the course (post-test = 25.10) was found to be significant for the sub-dimension of attitudes towards the measurement and evaluation dimension of the primary mathematics curriculum [$t(9) = -2.946$, $p < 0.05$].

For attitudes towards elementary mathematics curriculum (the whole scale), the difference between the mean score of the scale applied at the beginning of the course (pretest = 96.10) and the mean score of the scale applied at the end of the course (post-test = 103.10) was found to be significant [$t(9) = -7.826$, $p < 0.05$].

However, according to Green and Salkind (2016), more than the presence or absence of a difference between two mean scores in the t-test for related samples is needed. Because this finding needs to inform us about the difference's magnitude, it is recommended to test statistical significance and give the effect size. In this context, effect sizes are reported in Table 3.

**Table 3** Effect Sizes of Prospective Teachers on the MPDS Scale

| | s | Cid low | Cid upper | d |
|---|---|---|---|---|
| Attitudes towards the outcomes of primary mathematics curriculum | 1.87 | 1.974 | .338 | 1.174 |
| Attitudes towards the content of primary mathematics curriculum | 1.90 | 1.490 | .057 | .789 |
| Attitudes towards the learning-teaching processes of the primary mathematics curriculum | 1.08 | 2.255 | .486 | 1.389 |
| Attitudes towards the measurement and evaluation dimension of the primary mathematics curriculum | 1.93 | 1.665 | .164 | .932 |
| attitudes towards primary mathematics curriculum | 2.82 | 3.744 | 1.177 | 2.475 |

The effect size for attitudes towards the outcomes of the primary mathematics curriculum was calculated as d= 1.174. According to Cohen (2013) effect size classification, it is at a significant effect level. This shows that the "Comparison of Mathematics Curricula" course, based on comparing countries' mathematics curricula, significantly affects preservice mathematics teachers' attitudes towards the elementary mathematics curriculum.

The effect size for attitudes towards the content of the primary mathematics curriculum was calculated as d=.789. According to the effect size classification, it is at an average effect level (Cohen, 2013). This shows that the effect of the "Comparison of Mathematics Curricula" course, based on comparing countries' mathematics curricula, on preservice mathematics teachers' attitudes towards the elementary mathematics curriculum is at a medium level.

The effect size of the elementary mathematics curriculum on attitudes towards learning-teaching processes was calculated as d= 1.389. According to the effect size classification, it is at a significant effect level (Cohen, 2013). This shows that the "Comparison of Mathematics Curricula" course, based on comparing countries' mathematics curricula, significantly affects preservice mathematics teachers' attitudes towards the elementary mathematics curriculum.

The effect size for the attitudes towards the measurement and evaluation dimension of the primary mathematics curriculum was calculated as d=.932. According to the effect size classification, it is at a significant effect level (Cohen, 2013). This shows that the "Comparison of Mathematics Curricula" course, based on comparing the mathematics curricula of countries, significantly affects preservice mathematics teachers' attitudes towards the elementary mathematics curriculum.

The effect size for attitudes towards the primary mathematics curriculum was d= 2.475. According to the effect size classification, it is at a significant effect level (Cohen, 2013). This shows that the "Comparison of Mathematics Curricula" course, based on comparing countries' mathematics curricula, significantly affects preservice mathematics teachers' attitudes towards the elementary mathematics curriculum.

**Findings Related to Qualitative Measurements**

Within the scope of the research, the changes experienced by preservice mathematics teachers during the "Comparison of Mathematics Curricula" course emerged as a 4-stage process. The process is shown in Figure 1.

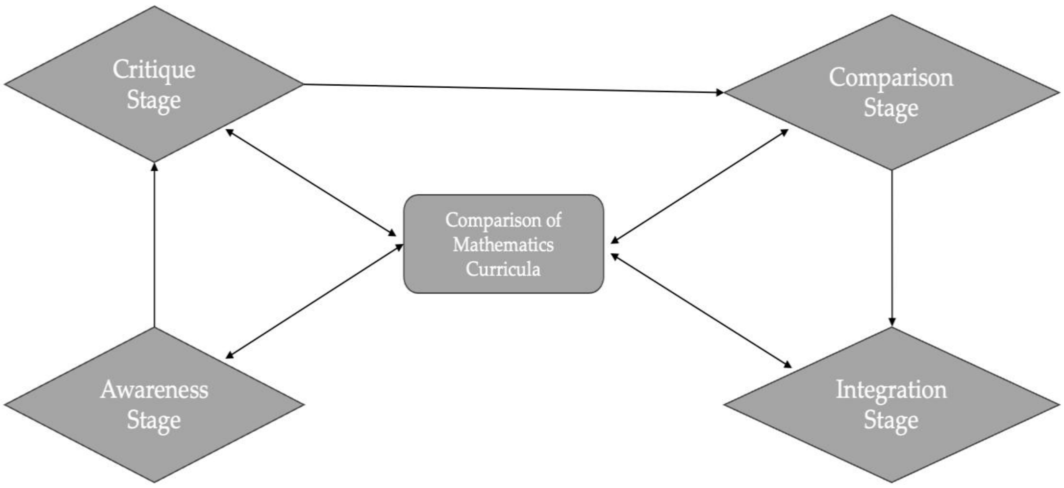



**Figure 1** 4-Stage Change Process

### *Phase 1: Awareness Phase*

The preservice teachers are in the process of becoming more familiar with and aware of their own country's mathematics curriculum. At this stage, they discovered the characteristics and structures of their mathematics curriculum. The codes of this category are given in Table 4.

**Table 4** Codes Related to Awareness Formed by Prospective Teachers' Examination of Mathematics Curriculum

| Category | Codes |
|---|---|
| Awareness | Objectives |
| | Acquisitions |
| | Structure of the programme |
| | Implementation |
| | Distribution of learning areas according to classes |
| | Measurement and evaluation |

It was understood that the participants needed to be aware of our primary mathematics curriculum before the implementation. It was seen that with the applied lesson, they gained ideas and information about the aims (f: 8), achievements (f: 9), structure (f: 8), implementation (f: 7), distribution of learning areas according to classes (f: 10), measurement and evaluation approaches (f: 8) of the primary mathematics curriculum in Türkiye. Sample participant opinions regarding this stage are as follows:

(K: 10): "Before the class, I had only heard the name of our mathematics programme, but the course made me understand the programme's purpose and what it aims to provide us." - A participant gaining awareness about its objectives.

(K: 5): "I used to know only the general framework of the programme; now I see how diverse and comprehensive the outcomes are." - Gaining detailed knowledge about the outcomes.

(K: 4): "I learned that assessment and evaluation techniques are not just limited to tests, but actually, a variety of more effective methods can be used to track students' development." - Gaining new perspectives on assessment and evaluation approaches.

### *Phase 2: Criticism Phase*

Preservice teachers critically evaluate the mathematics curriculum of their own country. They identified shortcomings, strengths, difficulties in implementation and areas for improvement. The codes of this category are given in Table 5.

**Table 5** Codes Related to the Criticism of Preservice Teachers by Examining Different Mathematics Curricula

| Category | Codes |
|---|---|
| Criticism | Weaknesses |
| | Has its strengths |
| | Should be simplified |
| | Repeated acquisitions |
| | Level |
| | Difficult in practice |
| | Ignored individual differences |

Participants started to express their criticisms by examining different mathematics curricula. In the course process, they criticised the Turkish elementary mathematics curriculum and the mathematics curricula of other countries together and emphasised that the curricula had weaknesses (f: 1), strengths (f: 9), needed to be simplified (f: 5), had repetitive acquisitions (f: 4), were level appropriate (f: 10), had difficulties in implementation (f: 2), and ignored individual differences (f: 4). Sample participant opinions regarding this stage are as follows:

(K: 1): "I noticed that some concepts are unnecessarily repeated across grades, which could be streamlined to make room for more diverse topics." - Addressing the issue of repetitive outcomes.

(K: 4): "The lack of consideration for individual differences in learning paces and styles is a significant oversight in these programmes. Each student's unique understanding should be addressed more explicitly." - Critiquing the programmes for overlooking individual differences.

(K: 6): "While the curriculum aims at a comprehensive education, the level of difficulty in some areas might not be suitable for all students, leading to potential challenges in classroom implementation." - Discussing level appropriateness and possible difficulties in application.

***Phase 3: Comparison Phase***

The preservice teachers examined different countries' mathematics curricula and compared them with their curricula. At this stage, they evaluated different approaches, teaching methods and objectives. The codes of this category are given in Table 6.

**Table 6** Codes Related to the Comparison Made by Preservice Teachers by Examining Different Mathematics Curricula

| Category | Codes |
|---|---|
| Comparison | Cultural differences reflected in teaching programmes<br>Learning areas show differences<br>Simplicity<br>Acquisitions are distributed to different classes<br>Different methods<br>Differences in objectives and approaches |

Participants made comparisons by examining different mathematics curricula. In the course process, they compared the Turkish elementary mathematics curriculum with the mathematics curricula of other countries and emphasised that cultural differences were reflected in the curricula (f: 8), differences in learning areas (f: 9), differences in terms of simplicity (f: 6), distribution of learning outcomes to different classes (f: 9), different methods (f: 10), differences in objectives and approaches (f: 7), and similarities (f: 10). Sample participant opinions regarding this stage are as follows:

(K:8): "Comparing the programme from Türkiye with those from other countries, it was exciting to see how each reflects its cultural values; for example, some programmes place a greater emphasis on collaboration and group work among students." - Reflecting on how cultural differences manifest in the programmes.

(K:2): "I noticed significant differences in the distribution of learning domains across different countries' mathematics programmes; some prioritise analytical thinking, while others highlight problem-solving skills." - Differences in learning domains.

(K:9): "Interestingly, the same outcomes are assigned to different programme grade levels. This might reflect varying understandings of when students are ready to learn these concepts." - Distribution of outcomes across different grades.

(K:10): "There are also big differences in the methods they use; some programmes rely more on visual and interactive tools, while others prefer traditional teaching methods." - Different methods contained within the programmes.

(K:1): "I observed that there are distinct differences in the objectives and approaches of each mathematics education programme; this diversity enriches the educational experience offered to students." - Differences in objectives and approaches.

### *Phase 4: Integration Phase*

Preservice teachers come to the stage of integrating the knowledge and perspectives gained during the course into their teaching practice and using them in practice. At this stage, they adapted the features of different curricula to their curricula and developed a more effective curriculum. The codes of this category are given in Table 7.

**Table 7** Codes Related to the Integration Realized by the Preservice Teachers' Examination of Different Mathematics Curricula

| Category | Codes |
|---|---|
| Integration | By international standards |
| | Simplified |
| | Different approaches and methods |
| | Student-centred |
| | By the measurement and evaluation approaches of international exams |

Participants examined different mathematics curricula and integrated them into their curriculum. During the course process, they recommended improving the Turkish primary mathematics curriculum. They offered integrated suggestions on increasing its compliance with international standards (f: 10), simplifying it (f: 10), including different approaches and methods (f: 10), increasing student-centeredness (f: 10), and making it suitable for the assessment and evaluation approaches of international exams (f: 10). Sample participant opinions regarding this stage are as follows:

(K:6): "I believe our programme should better align with international standards; this way, our students can gain a global perspective and become competitive worldwide." - Integration of recommendations to increase compliance with international standards.

(K:7): "I noticed that our mathematics education programme is overly complex and could be simplified to facilitate easier understanding for students." - Simplification of the programme.

(K:2): "We've realised the need to integrate various approaches and methods into our programme to cater to different learning styles. This can make it easier for each student to learn math in their own way." - Integration of recommendations to include different approaches and methods.

(K:5): "I want to emphasise the importance of a student-centred approach; having more interactive lessons and encouraging student participation can enhance their learning process." - Increase student-centeredness.

(K:9): "Aligning our mathematics programme with the assessment approaches of international exams could help our students perform better in these exams." - Integration of recommendations to align with the assessment approaches of international exams.

## Results, Conclusions and Suggestions

This study, conducted with a one-group pretest-posttest experimental design, was supported by interviews and aimed to find more comprehensive and in-depth answers to the research questions. The results provide essential insights for evaluating and developing the primary mathematics curriculum.

Quantitative findings show that the "Comparison of Mathematics Curricula" course significantly impacted preservice mathematics teachers' attitudes towards the elementary mathematics curriculum. When these effects are examined in terms of the curriculum's objectives, content, learning-teaching processes and assessment and evaluation dimensions, it is understood that the course has a comprehensive effect on mathematics teaching. In particular, the high effect sizes of the curriculum on learning-teaching processes and general attitude indicate that significant improvements and changes can be made in these areas.

On the other hand, the interviews' findings reveal the pre-service teachers' in-depth views and experiences about the primary mathematics curricula. The participants' suggestions on how the curricula should better comply with international standards, be simplified, and include a variety of teaching methods and approaches indicate that the current curricula need to be improved. These suggestions emphasize the need to increase student-centeredness in mathematics teaching and to adopt assessment and evaluation approaches suitable for international exams.

The results show the significant impact of primary mathematics curricula on pre-service teachers' attitudes and how the current state of the curricula and potential areas for improvement are perceived from the participants' perspective. This bidirectional perspective provides a solid basis for maintaining the strengths of the programmes and making structured changes in areas that need improvement. Therefore, this study makes valuable contributions to a more effective and inclusive revision of primary mathematics curricula in terms of educational policies and practices.

The quantitative findings obtained in this study indicate that the course "Comparison of Mathematics Teaching Programmes" significantly impacts teacher candidates' attitudes towards primary school mathematics teaching programmes. The substantial effect size of attitudes towards the teaching-learning processes and the overall programme is particularly noteworthy. This finding is consistent with previous research; for example, Tosuncu (2019) noted the significant effects of mathematics teaching programmes on the pedagogical content knowledge of teacher candidates. Additionally, these findings are in line with the emphasis in TIMSS (2019) and PISA (2018) reports on the critical role of programme quality and content on student achievement (Mullis et al., 2020; OECD, 2019). Furthermore, teachers' perspectives on curriculum are critical in implementing programmes and student achievement. A study by Swars et al. (2009) showed that teachers' positive attitudes towards mathematics teaching can significantly enhance students' mathematics achievement. These findings are consistent with the results obtained regarding the importance of teacher candidates' attitudes towards primary school mathematics teaching programmes in this study. Teachers' positive attitudes towards programmes can contribute to their more effective implementation and thus increase student achievement (Swars et al., 2009).

Moreover, Di Martino and Zan's (2010) research indicates that students' positive attitudes towards mathematics lessons positively affect mathematics achievement. The findings of this study suggest that recommendations such as simplifying curricula and increasing student-centeredness can improve students' attitudes towards mathematics learning. Adopting student-centred approaches can contribute to students developing positive attitudes towards mathematics learning, thereby enhancing their success (Di Martino & Zan, 2010).

The results of international assessment programmes such as TIMSS and PISA demonstrate that the quality of curriculum in mathematics education determines student achievement. For instance, countries like Singapore and Finland consistently achieve high levels of success due to their student-centred learning approaches and emphasis on critical thinking skills (OECD, 2019; Mullis et al., 2020). The findings of this study suggest that similar improvements in primary school mathematics teaching programmes in Türkiye could better align with international standards and enhance student achievements. Moreover, international assessment programmes like TIMSS and PISA reveal the global differences in mathematics education and their impact on student achievement. Particularly, students in Asian countries demonstrate high levels of achievement in mathematics, some of which can be attributed to the attitudes of students and teachers towards mathematics teaching programmes (Cai, 2004; Leung, 2002).

The findings obtained from the interviews provide suggestions for increasing the alignment of pre-service teachers with international standards, simplifying the curriculum content and strengthening student-centredness in primary school mathematics teaching programmes. These suggestions align with the standards and recommendations published by NCTM (National Council of Teachers of Mathematics, 2020). NCTM emphasises the importance of student-centredness and problem-solving skills in mathematics education and states that the curricula of countries that perform well in international assessments have these characteristics.

Additionally, examining the mathematics curricula of other countries reveals different approaches and emphases. For instance, countries like Finland and Singapore achieve high success levels by emphasising student-centred learning approaches and critical thinking skills (Finnish National Agency for Education, 2014; Ministry of Education, Singapore, 2020). These countries also effectively use technology integration in education (Clark-Wilson et al., 2020). In countries like China, it is observed that focusing on mathematical problems from an early age and the beliefs of teachers are essential (Cai, 2004). These approaches suggest that countries like Türkiye should review their curricula and improve by examining international successes (Atweh & Clarkson, 2002; Tatto & Senk, 2011).

Furthermore, prospective teachers' examination of international mathematics curricula provides them with knowledge about various teaching methods and content, thereby presenting significant opportunities to improve their curricula further (Cai et al., 2017; Kaiser, 2002). Such comparative studies can also enhance educational collaborations and exchanges between countries, thus promoting the development of broader and more inclusive educational programmes (Zhou et al., 2020).

The findings of this study emphasise the need for the development of primary school mathematics curricula and the essential contributions that pre-service teachers can make in this process. Moreover, it is highlighted that curriculum alignment with international standards, student-centredness and diversity can increase students' mathematics achievement. Therefore, it is suggested that educational policies and practices should be shaped according to these findings. Attitudes towards mathematics curricula have essential effects at both teacher and student levels, and these attitudes can directly affect student achievement. Positive attitudes of teachers and students can contribute to more effective implementation of mathematics curricula and increase student achievement. Based on the research findings, the following suggestions can be made:

Develop education programmes that will improve teacher candidates' attitudes towards mathematics teaching programmes and equip them with knowledge about various teaching methods. These programmes can inform teacher candidates about international mathematics teaching programmes and approaches, enabling them to diversify teaching methods and create student-centred learning environments.

Türkiye's current primary school mathematics teaching programmes should be reviewed and updated in light of international standards and indicators of success regarding their content, teaching-learning processes, and assessment approaches. The importance of student-centeredness and practical applications should be emphasised throughout this process.

Education collaboration should be established with countries ranking high in international achievement rankings. Teacher and student exchange programmes should be encouraged to facilitate the exchange of knowledge about different teaching approaches and educational systems.

Programmes for the continuous professional development of mathematics teachers should be created and promoted. These programmes should help teachers enhance their skills in utilising new teaching methods, assessment approaches, and technological tools.

Research on mathematics teaching and curriculum should be encouraged, and necessary resources should be provided to develop innovative approaches in this field. Integrating research findings into teaching practice will enhance the quality of mathematics education.

**Compliance with Ethical Standards**

*Disclosure of potential conflicts of interest*

No conflict of interest.

*Funding*

None.

*CRediT author statement*

The study was authored, and the whole process was carried out by the corresponding author.

*Research involving Human Participants and/or Animals*

The study involves human participants. Ethics committee permission (Date: 10.11.2023, Number: 11- 2023-524) was obtained from Necmettin Erbakan University, Social and Human Sciences Research Ethics Committee.

## Uluslararası Matematik Öğretim Programlarını Karşılaştırma Deneyimlerinin Matematik Öğretmen Adaylarının İlköğretim Matematik Programına İlişkin Görüşlerine Etkisi

**Özet:**

Bu araştırma, 2023-2024 eğitim-öğretim yılında Türkiye'nin İç Anadolu Bölgesi'nde bir üniversitede matematik öğretmeni adayları arasında gerçekleştirilmiştir. Araştırmanın temel amacı, uluslararası matematik ders programlarının karşılaştırılması deneyimlerinin, matematik öğretmen adaylarının Türkiye'deki ilköğretim matematik öğretim programına ilişkin görüşlerine etkisini incelemektir. Özellikle, matematik öğretmen adaylarının uluslararası matematik ders programlarını inceleyerek elde ettikleri deneyimlerin, kendi ülkelerinin ilköğretim matematik öğretim programına yönelik algılarına ve tutumlarına nasıl yansıdığı araştırılmıştır. Araştırma, tek gruplu ön test son test bir yöntem kullanılarak tasarlanmıştır; nicel veriler görüşmelerle desteklenmiş ve analiz edilmiştir. Örneklem, "Matematik Öğretim Programlarının Karşılaştırılması" seçmeli dersini alan 10 matematik öğretmeni adayından oluşmaktadır. Bulgular, katılımcıların uluslararası matematik öğretim programlarını inceleme deneyimlerinin, öğretim programlarına yönelik algılarını ve tutumlarını olumlu yönde etkilediğini göstermektedir. Öğretmen adayları, uluslararası standartlara daha yakın bir program geliştirilmesi, öğrenci merkezliliğin artırılması, ölçme ve değerlendirme yöntemlerinin çeşitlendirilmesi gibi konularda önerilerde bulunmuştur. Araştırma sonuçları, ilköğretim matematik öğretim programlarının geliştirilmesi için önemli iç görüler sağlamaktadır ve öğretmen eğitimi programlarında uluslararası perspektiflerin entegre edilmesinin önemini vurgulamaktadır.

*Anahtar kelimeler*: Matematik öğretim programı, öğretmen adayları, deneysel desen.

## References

Atteh, E., Assan-Donkoh, I., Ayiku, F., Nkansah, E., & Adams, A. K. (2020). The use of technology among school mathematics teachers and students: The new wave of recommended instructions. *Asian Research Journal of Mathematics*, *16*(5), 18–29. https://doi.org/10.9734/arjom/2020/v16i530189

Atweh, B., & Clarkson, P. (2002). Globalized curriculum or global approach to curriculum reform in mathematics education. *Asia Pacific Education Review, 3*(2), 160–167. https://doi.org/10.1007/bf03024909

Büyüköztürk, Ş., Kılıç Çakmak, E., Akgün, Ö. E., Karadeniz, Ş., & Demirel, F. (2011). *Bilimsel araştırma yöntemleri*. Pegem.

Cai, J. (2004). Why do U.S. and Chinese students think differently in mathematical problem solving? Influence of early algebra learning and teachers' beliefs. *Journal for Research in Mathematics Education*, *35*(5), 297–328. https://doi.org/10.1016/j.jmathb.2004.03.004

Cai, J., Mok, I., Reddy, V., & Stacey, K. (2017). International comparative studies in mathematics: Lessons and future directions for improving students' learning. In G. Kaiser (Ed.), *Proceedings of the 13th International Congress on Mathematical Education* (pp. 79–99). Springer International Publishing. https://doi.org/10.1007/978-3-319-62597-3_6

Clark-Wilson, A., Robutti, O., & Thomas, M. (2020). Teaching with digital technology. *ZDM – Mathematics Education, 52*(7), 1223-1242. https://doi.org/10.1007/s11858-020-01196-0

Cohen, J. (2013). *Statistical power analysis for the behavioral sciences*. Academic Press.

Creswell, J. W. (2018). *Nitel araştırma yöntemleri: Beş yaklaşıma göre nitel araştırma ve araştırma deseni (*Bütün, M., & Demir, S. B. Çev). Siyasal.

Di Martino, P., & Zan, R. (2010). 'Me and maths': Towards a definition of attitude grounded on students' narratives. *Journal of Mathematics Teacher Education, 13*, 27-48. https://doi.org/10.1007/s10857-009-9134-z

Finnish National Agency for Education (2014). "National Core Curriculum for Basic Education." Available at https://www.oph.fi/en/statistics-and-publications/publications/national-core-curriculum-basic-education-2014

Green, S. B., & Salkind, N. J. (2016). *Using SPSS for Windows and Macintosh books a la carte*. Pearson.

Hjalmarson, M. A. (2008). Mathematics curriculum systems: Models for analysis of curricular ınnovation and development. *Peabody Journal of Education*, *83*(4), 592–610. https://doi.org/10.1080/01619560802414965

Kaiser, G., Leung, F. K.S., Romberg, T., & Yaschenko, I. (2002). International comparison in mathematics education: An overview. *ICM, 1*, 631-646. https://arxiv.org/pdf/math/0212416v1

Kultusministerkonferenz (KMK) (2020). Educational standards in mathematics for the primary certificate. Germany. https://www.kmk.org/fileadmin/

Leung, F. K. S. (2002). Behind the high achievement of East Asian students. *Educational Research and Evaluation*, *8*(1), 87-108. https://doi.org/10.1076/edre.8.1.87.6920

Mills, G. E., & Gay, L. R. (2016). *Educational research: Competencies for analysis and applications*. Pearson.

Ministry of Education, Culture, Sports, Science and Technology, Japan. (2020). Course of study for elementary school. http://www.mext.go.jp/en/policy/education/

Ministry of Education, Singapore. (2020). Mathematics syllabus: Primary. https://www.moe.gov.sg/education/syllabuses/sciences/

Mullis, Ina V.S., Martin, Michael O., Foy, Pierre, Kelly, Dana L., Fishbein, Bethany (2020). *TIMSS 2019 international results in mathematics and science*. TIMSS & PIRLS International Study Center, Boston College. https://www.iea.nl/sites/default/files/2020-12/TIMSS%202019-International-Results-in-Mathematics-and-Science.pdf.

National Council of Teachers of Mathematics. (2020). NCTM Standards and Guidelines for Mathematics Teaching. https://www.nctm.org/Standards-and-Positions/CAEP-Standards/

Nguyen, T. T., Trinh, T. P. T., Ngo, H. T. V., Hoang, N. A., Tran, T., Pham, H. H., & Bui, V. N. (2020). Realistic mathematics education in Vietnam: Recent policies and practices. *International Journal of Education and Practice, 8*(1), 57-71. https://doi.org/10.18488/journal.61.2020.81.57.71

OECD. (2019). PISA 2018 results (Volume I): What students know and can do? https://www.oecd.org/publications/pisa-2018-results-volume-i-5f07c754-en.htm

Patton, M. Q. (2014). *Nitel araştırma ve değerlendirme yöntemleri.* (Çev. Ed. M. Bütün, S.B. Demir). Pegem Akademi.

Peter-Koop, A., & Scherer, P. (2012). Early childhood mathematics teaching and learning. *Journal für Mathematik-Didaktik*, *33*, 175-179. https://doi.org/10.1007/s13138-012-0043-9

South Korean Education and Research Information Service (KERIS). (2020). Mathematics curriculum for elementary school. https://english.keris.or.kr/

Swars, S. L., Smith, S. Z., Smith, M. E., & Hart, L. C. (2009). A longitudinal study of effects of a developmental teacher preparation program on elementary prospective teachers' mathematics beliefs. *Journal of Mathematics Teacher Education, 12,* 47-66. https://doi.org/10.1007/s10857-008-9092-x

Tatto, M T., & Senk, S L. (2011). The mathematics education of future primary and secondary teachers: Methods and findings from the teacher education and development study in mathematics. *Journal of Teacher Education, 62*(2), 121-137. https://doi.org/10.1177/0022487110391807

Tosuncu, A. (2019). *Investigation of the factors affecting the pedagogical content knowledge of prospective elementary mathematics teachers using hierarchical linear modelling* (Publication No. 549435) [Master's thesis, Karamanoğlu Mehmetbey University]. Council of Higher Education Thesis Center.

US Department of Education. (2010). Common core state standards for mathematics. http://www.corestandards.org/Math/

Wößmann, L. (2003). Schooling resources, educational institutions and student performance: the international evidence. *Oxford Bulletin of Economics And Statistics*, *65*(2), 117-170. https://doi.org/10.1111/1468-0084.00045

Yıldız, A., & Gürgen, L. (2021). Evaluation of primary school mathematics curriculum according to the expertise-oriented approach to curriculum evaluation. *Turkish Studies-Educational Sciences*, *16*(6), 2717-2734. http://dx.doi.org/10.7827/TurkishStudies.51910

Zhang, N., Yang, R., Liu, Y., Zhou, J., & Wang, G. (2018). Integration of information technology (IT) and the mathematics curriculum. In Y. Cao & F. K. S. Leung (Eds.), *The 21st century mathematics education in China* (pp. 259-278). https://doi.org/10.1007/978-3-662-55781-5_13

Zhou, G., Wang, P., Tian, L., Zhang, J., Li, Y., Fu, C., & Wu, S. (2020). What they learned won't go away: The impacts of an international exchange program on Chinese teacher candidates' understanding of and practice in science education. *ECNU Review of Education*, *5*(3), 404-424. https://doi.org/10.1177/2096531120959676

Necatibey Faculty of Education Electronic Journal of Science and Mathematics Education
Vol. 18, Issue 2, December 2024, pp. 271-290. ISSN: 1307-6086

Research Article

# A Bibliometric Analysis of Laboratory Safety: Its Significance for the Discipline of Chemistry*

**Mustafa Tuğbay EMİROĞLU[1], Ayhan YILMAZ[2]**

[1] Hacettepe University, Faculty of Education, Türkiye, tugbayemiroglu@hacettepe.edu.tr, https://orcid.org/0009-0000-6926-060X

[2] Hacettepe University, Faculty of Education, Türkiye, ayhany@hacettepe.edu.tr, http://orcid.org/0000-0003-4252-5510

Received : 19.07.2024 Accepted : 19.08.2024

Doi: https://doi.org/10.17522/balikesirnef.1519282

---

*Abstract* – Laboratory safety is a multidisciplinary field of research that is highly significant for science education, and especially for the discipline of chemistry. This study addresses the developments, primary fields of study, and research tendencies related to the scope and application of laboratory safety in chemistry. The study utilized a survey research design from among descriptive methods and applied a bibliometric analysis technique. To this end, a total of 279 works published on laboratory safety between 1965 and 2024 and indexed in the Web of Science database were analyzed. The analyses encompassed 60 countries, 279 publications, 51 authors, and 43 subject categories. Findings were analyzed using VOSviewer via the bibliometric analysis method. According to the obtained results, the country with the highest number of relevant publications is the United States, the institution with the highest number of published articles is the University of California system, the journal with the highest number of articles is Journal of Chemical Education, and the researchers with the highest numbers of publications are Nancy L. Wayne and Xiaoyan Wang, followed by Imke Schröder, James H. Gibson, and Xinglong Jin. Although many researchers study the subject of laboratory safety, it is seen that the cooperation among authors is extremely weak. The most commonly used keywords were identified as “laboratory safety,” “safety hazards,” and “laboratory management.”

*Keywords*: Bibliometric mapping, VOSviewer, laboratory safety, chemistry.

------------

Corresponding author: Mustafa Tuğbay EMİROĞLU, tugbayemiroglu@hacettepe.edu.tr

*This article was produced from data obtained within the scope of the corresponding author’s doctoral dissertation.

## Introduction

Laboratories, as environments of teaching and learning, are a fundamental part of science, technology, and engineering education at all levels. As chemistry is a branch of science based particularly heavily on laboratory research, emphasis must always be placed on laboratory safety in chemistry education (Hill, 2007).

While laboratories are the most effective spaces for putting theoretical knowledge into practice, they are also a complementary part of the discipline of chemistry. To reach the goals set for chemistry education, it is necessary to use laboratories effectively. Laboratory applications help students meaningfully and permanently structure the obtained knowledge, improve their psychomotor skills, use their hands effectively, and enhance their communication skills (Atasoy, 2002; Çepni et al., 1994; Hofstein & Lunetta, 2004; Morgil & Yılmaz, 2000). Previous studies have emphasized that laboratory work is significantly effective in the scientific success of students and helps students develop positive attitudes toward science (Hofstein et al., 2005).

A laboratory learning environment is a complicated learning environment. In this context, successful learning requires scientific reasoning, creativity and problem-solving, use of the affective domain, an understanding of the nature of science, meaningful and permanent structuring of obtained information, development of psychomotor skills, and unification and synchronization of the components of communication skills (Seery et al., 2018; Yılmaz, 2023).

### Laboratory Safety

In all applications conducted in a laboratory environment, the most important point to keep in mind is safety. Safety measures taken during laboratory applications should not restrict the practical studies that are being undertaken; on the contrary, their purpose is to help conduct those studies safely (Morgil & Yılmaz, 2000; Yılmaz, 2015).

The US Chemical Safety and Hazard Investigation Board reported over 120 accidents occurring in university laboratories in the United States in the decade between 2001 and 2011, resulting in many injuries along with millions of dollars of damage. Other studies have revealed the occurrence of many high-profile accidents in laboratories, resulting in serious injuries and even death (Menard et al., 2020; Zhu et al., 2018). In Türkiye, the major reason of the laboratory accidents is the failure to apply the necessary safety measures (Akpullukcu, 2017; Tepe & Tekbıyık, 2019).

When the international literature is reviewed in terms of laboratory safety, we can see that subjects such as laboratory safety education (Meyer, 2017; Sigmann, 2018), laboratory safety culture (Ayi & Hon, 2018; Walters et al., 2017; Yoon, 2021), laboratory safety management (Weil, 2016; Zhu et al., 2018), laboratory risk assessment (Omidvari et al., 2015; Pluess et al., 2016), safety information (Agustian & Seery, 2017; Miller et al., 2000), laboratory hazard symbols (Wangdi & Tshomo, 2016), and chemical laboratory safety awareness, attitudes, and practices (Ayana et al., 2017) have been explored.

**Bibliometric Analysis**

This study used bibliometric analysis to provide a general view of the pioneering researchers, most prolific countries and institutions, and most commonly researched topics in the field of laboratory safety. Previous studies in the literature have highlighted the fact that bibliometric analysis is suitable for use in scientific research and that it helps scholars review the literature in a very short period of time (Block & Fisch, 2020; De Bellis, 2009; Gutierrez-Salcedo et al., 2017; Kurutkan & Orhan, 2018; McBurney & Novak, 2002).

**Significance and Aim of the Study**

Since laboratory safety is a vital and multidisciplinary field of research, studies on laboratory safety have rapidly increased over the last twenty years. However, in Türkiye, due to the low number of studies on laboratory safety, various accidents occur, especially in laboratory applications in middle school science classes, and such accidents may pose serious risks (Akpullukcu, 2017; Demir, 2016; Şener, 2018; Morgil & Yılmaz, 2000; Yılmaz, 2015). In 34 recorded laboratory or experiment accidents in Türkiye between 2001 and 2017, some serious incidents resulted in the loss of sight, burns, injuries, or poisoning (Tepe & Tekbıyık, 2019). In the studies conducted by Stuart and Toreki (2014) and Olewski and Snakard (2017), a total of 533 accidents were documented in three years, considering accidents occurring in school laboratories, and information on their severity was confirmed with relevant statistics. Among individuals working in laboratories of universities and other schools, 45% reported having an accident, and 73.7% of those cases were reported to be due to chemical substances (Nasrallah et al., 2022).

There is one previous article on the bibliometric analysis of laboratory safety in the international literature, published by Yang et al. (2019). However, no such articles on this topic have been published in the Turkish literature to date. Thus, this study was undertaken to reveal the development of research conducted on laboratory safety; the most influential and prolific authors, countries, and institutions; and trends in international cooperation. Safety is of the utmost priority in laboratory work. Accordingly, evaluating specific topics that require special attention in the field of laboratory safety with a holistic approach will shed light on avenues for future studies and is expected to contribute to both the overall body of literature and the individual work of researchers in this field.

### Research Question

What is the bibliometric status of scientific studies conducted in the field of laboratory safety?

## Method

### Research Model

This study used a survey research design from among descriptive research methods. Bibliometric mapping was performed using VOSviewer and scientific studies in the field of laboratory safety were analyzed using the bibliometric analysis method. VOSviewer is an open-source bibliometric mapping program (www.neesjanvaneck.nl/vos/ ), and the possible analyses that can be performed with this software are shown in Table 1 (as cited by Arslan, 2022; Van Eck & Waltman, 2010, p. 536).

### Data Collection

This study used the Web of Science (WoS) database. Specifically, data were obtained from the WoS Core Collection Database ( https://clarivate.com/academia-government/scientific-and-academic-research/research-discovery-and-referencing/web-of-science/web-of-science-core-collection/ ), providing access to SCI-Expanded, SSCI, and ESCI citations. “Laboratory safety” and “safety in chemical laboratories” were used as keywords while searching the database. We found 1064 publications with the search phrase “laboratory safety” and 2945 publications with “safety in chemical laboratories.” The obtained publications were then reviewed in line with the inclusion/exclusion (filtering) criteria and the data pool for this study was created. All data were recorded on June 15, 2024, and a total of 279 publications were selected for inclusion in the study. Between 1965 and 2024, a total of 205 journal articles, 45 conference proceedings, 19 review articles, 8 meeting abstracts, and 2 early-access publications were obtained from different disciplines included the WoS categories. In the process of determining whether or not the publications found in the dataset were relevant within the scope of laboratory safety and its significance for the discipline of chemistry, the researchers reviewed the abstracts of all obtained publications.

Expert opinions were also collected regarding the publications obtained from the dataset and the data pool of the study was finalized in this way. The data collection process of the study is illustrated in Figure 1.

**Data sources:**
SCI-Expanded, SSCI, and ESCI citation index databases in the WoS Core Collection

**Topics searched:**
"Laboratory safety" and
"safety in chemical laboratories"
from 1965 to 2024

**Filtering:**
"Laboratory safety," 1064 documents
"Safety in chemical laboratories," 2945 documents
After filtering: 279 documents

WoS Core Collection on June 15, 2024

**Figure 1** Data Collection Process

**Data Analysis**

The survey research design from among descriptive research methods was used to examine the data obtained in this study and a bibliometric analysis technique was applied. The distribution of the 279 publications obtained in the data collection process was analyzed by years, authors, journals, institutions, and countries. Bibliometric analysis was initially developed by Van Eck & Waltman (2010, as cited by Arslan, 2022). Science mapping and performance analyses, also known as citation analyses, were done using VOSviewer software. The analysis types and analysis units utilized with VOSviewer are explained in Table 1.

**Table 1** Types and Units of Analysis Available with VOSviewer Mapping

| Type of analysis | Unit of analysis |
|---|---|
| Bibliographic-coupling | Documents, sources, authors, organizations, countries |
| Co-citation | Cited references, cited sources, cited authors |
| Co-authorship | Authors, organizations, countries |
| Co-occurrence | Author keywords |
| Citation | Documents, sources, authors, organizations, countries |

In the present study, co-citation, co-authorship, co-occurrence, and citation analyses were performed.

## Findings

The findings of this study are categorized into two sections: descriptive findings and bibliographic findings.

### Descriptive Findings

Concerning laboratory safety in the discipline of chemistry, the distribution of the 279 relevant publications obtained from the WoS database up to June 2024 was evaluated based on years, on authors with the highest numbers of publications, on journals with the highest numbers of articles, and on countries and institutions with the highest numbers of publications.

The distribution of articles according to years is presented in Figure 2. Considering the distribution of articles by years, it is seen that 10 articles between 1965 and 1980, 13 articles between 1988 and 1999, and 10 articles between 2002 and 2007 were published on laboratory safety in chemistry education. Furthermore, there was no increase in the number of articles published until 2009, but there were 7 articles published in 2010 and the highest numbers of articles were published between 2021 and 2023. The increase in the number of publications is an indicator of the importance of the field of laboratory safety at every level of education.

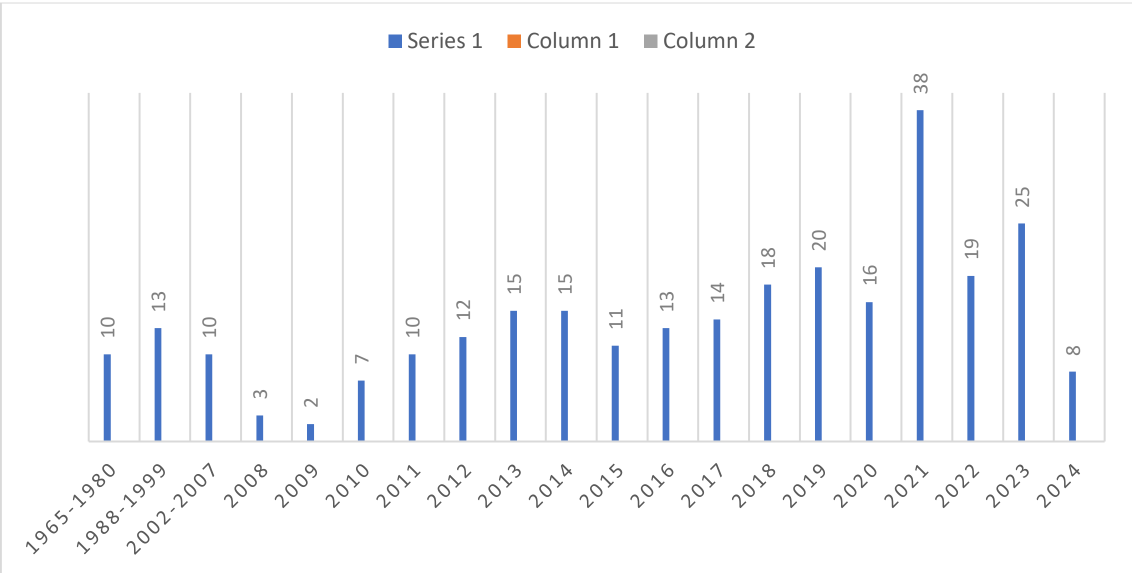



**Figure 2** Yearly Distribution (1965-2024)

The names of the five researchers who published the highest numbers of articles in this field are given in Table 2. The researcher with the most publications in this field was Nancy L. Wayne, followed by Xiaoyan Wang, Imke Schröder, James H. Gibson, and Xinglong Jin.

**Table 2** The Five Researchers with the Highest Numbers of Articles

| Author's name | Number of article |
|---|---|
| Nancy L. Wayne | 6 |
| Xiaoyan Wang | 6 |
| Imke Schröeder | 5 |
| James H. Gibson | 5 |
| Xinglong Jin | 5 |

The names of the eight journals with the highest numbers of articles in this field are given in Table 3. The five journals that published the highest numbers of relevant articles were found to be Journal of Chemical Education, Journal of Chemical Health and Safety, ACT Chemical Health Safety, Abstracts of Papers of the American Chemical Society, and Safety Science.

**Table 3** The Eight Journals with the Highest Numbers of Articles

| Journal | Number of article |
|---|---|
| Journal of Chemical Education | 52 |
| Journal of Chemical Health Safety | 37 |
| ACT Chemical Health Safety | 23 |
| Abstracts of Papers of the American Chemical Society | 8 |
| Safety Science | 7 |
| Process Safety Progress | 7 |
| Journal of Loss Prevention in the Process Industries | 6 |
| ACS Symposium Series | 5 |

The countries with the highest numbers of publications in this field of research are given in Table 4. The remaining countries (39 countries with a total of 41 articles) were all countries with 3 or fewer relevant publications. Considering the distribution of publications based on countries, the United States had the most, followed by China, Brazil, England, and South Korea.

**Table 4** Countries with the Highest Numbers of Publications

| Country | Number of article | Country | Number of article |
|---|---|---|---|
| USA | 104 | Germany | 6 |
| China | 35 | Spain | 6 |
| Brazil | 11 | Switzerland | 6 |
| Canada | 10 | Taiwan | 5 |
| England | 8 | Thailand | 5 |
| South Korea | 8 | Slovakia | 4 |
| France | 7 | India | 4 |
| Malaysia | 7 | Other countries (n=39) | 41 |
| Türkiye | 7 | | |

The names of the 13 institutions with the highest numbers of publications are given in Table 5. The institution with the most relevant publications was the University of California system, followed by University of California-Los Angeles, the American Chemical Society, Ecole Polytechnique Federale de Lausanne, and Swiss Federal Institutes of Technology Domain.

**Table 5** The 13 Institutions with the Numbers of Articles

| Institutions | Number of article |
|---|---|
| University of California system | 10 |
| University of California-Los Angeles | 8 |
| American Chemical Society | 5 |
| Ecole Polytechnique Federale de Lausanne | 5 |
| Swiss Federal Institutes of Technology Domain | 5 |
| Tianjin University of Technology | 5 |
| University of North Carolina | 5 |
| Battelle Memorial Institute | 4 |
| United States Department of Energy (DOE) | 4 |
| University System of Ohio | 4 |
| KU Leuven | 3 |
| Wittenberg University | 2 |
| Abant Izzet Baysal University | 1 |

## Bibliographic Findings

### *Co-authorship results*

In the analysis of co-authorship in the field of laboratory safety in chemistry education, a network map was created to identify the authors with the most links, representing the most cooperation (Minimum number of publications of an author: 1; Minimum number of citations of an author: 1; Number of authors to be selected: 630; Clusters: 3; Links: 49; Total link strength: 61). In Figure 3, three different colors and types of nodes are apparent. Each color represents clusters of authors who conduct research on similar topics. On the other hand, the lines represent cooperation between authors. Visualizing the relationship between authorship and co-authorship in the field of laboratory safety helps reveal the most prolific authors. The most prolific authors were accordingly found to be Nancy L. Wayne (6 publications), James H. Gibson (5 publications), and Imke Schröder (4 publications), who are shown with blue nodes in Figure 3 together with their cooperation with other authors.

Although many scientists have studied laboratory safety, the cooperation among authors in this field is notably weak.

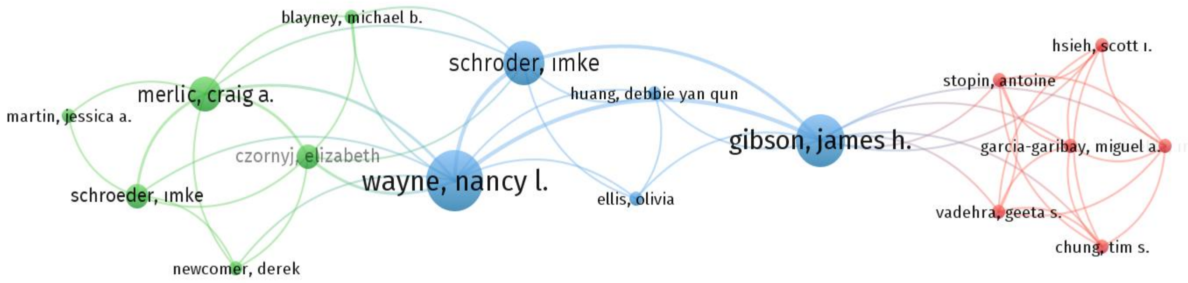



**Figure 3** Network Visualization Map of Co-Authorship

### *Citations of authors*

To visualize the network of citations in this field, a network map was created (Minimum number of documents of an author: 1; Minimum number of citations of an author: 1; of 775 authors, 582 met these thresholds; Clusters: 15; Links: 2908; Total link strength: 3310). Generally speaking, the most prolific authors and well-known research groups contribute greatly to science and to the work of other researchers, and this is also reflected in the numbers of citations that their publications receive. Therefore, the citation map given in Figure 4 was created. The most cited authors were found to be James H. Gibson (160 citations), Nancy L. Wayne (143 citations), Imke Schröder (121 citations), and A.D. Menard and J.F. Trant (108 citations).

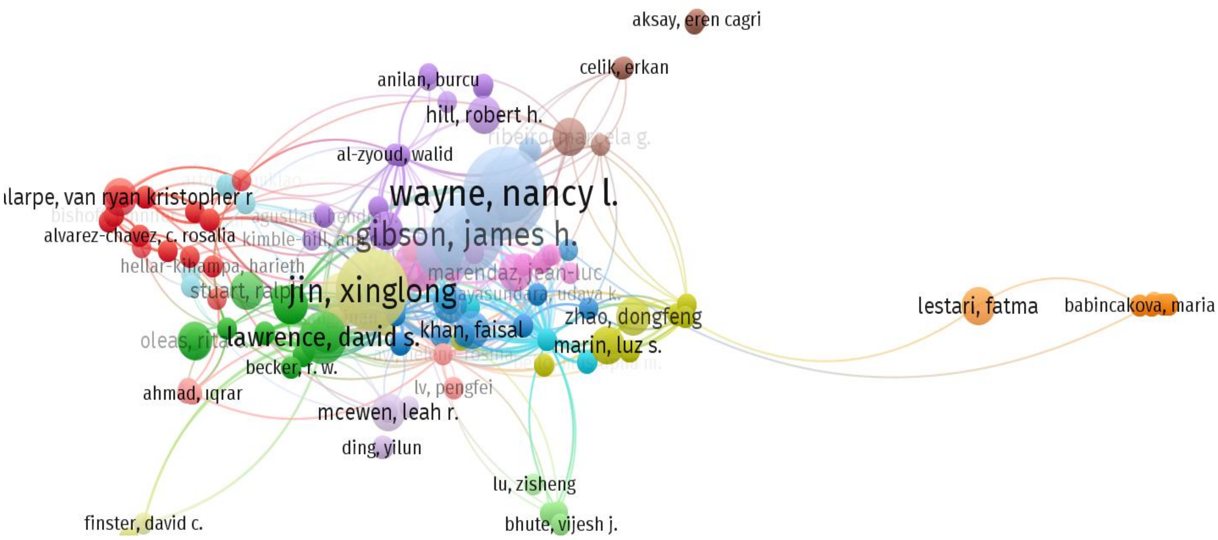



**Figure 4** Network Visualization Map of Citations

**Table 6** The Five Most Frequently Cited Documents on Laboratory Safety

| Article | Journal | Country | Citation | Average per year |
|---|---|---|---|---|
| Menard, A.D. and Trant, JF A (2020) review and critique of academic lab. safety research | Nature Chemistry | Canada | 109 | 21.8 |
| Agustian, H. Y and Seery, M.K (2017). Reasserting the role of pre-laboratory activities in chemistry education: a proposed framework for their design | Chemistry Education Research And Practice | England | 90 | 11.25 |
| Yang, YF , Reniers, G , Chen, Guohua and Goerlandt, F (2019). A bibliometric review of laboratory safety in universities | Safety Science | China Belgium Netherlands Canada | 77 | 12.83 |
| Schröder, I, Huang, DYQ Ellis, O, Gibson, JH, Wayne, NL.(2016) Laboratory safety attitudes and practices: A comparison of academic, government, and industry researchers | Journal Of Chemical Health & Safety | USA | 75 | 8.33 |
| Walters, AUC; Lawrence, W; Jalsa, NK (2017). Chemical laboratory safety awareness, attitudes and practices of tertiary students | Safety Science | Univ. West Indie | 64 | 8 |

Citation analysis is crucial in determining the effect and quality of a published article. Looking at Table 6, the most frequently cited publication is “A review and critique of academic lab safety research” by Menard and Trant (2020). This article also has the highest annual number of citations. In this work, Menard and Trant (2020) revealed that many high-profile accidents resulting in serious injuries and death in academic laboratories around the world occurred in the last decade. Furthermore, they suggested that laboratory safety policies have still not been improved even after such incidents, and they defined the research questions to be asked to minimize serious academic laboratory accidents in the future and highlighted the need for determined leadership in this area. This article has been cited 109 times to date.

### *Co-citations of references*

The co-citation map of reference analysis is given in Figure 5. This analysis was conducted to evaluate the interaction among studies done on laboratory safety and the publications with co-citations (Minimum number of citations of cited references: 3; of 6295 cited references, 204 met the threshold; Clusters: 7; Links: 6364).

Nodes of the same color represent the same cluster comprising articles on similar topics. The size of each node signifies the number of citations an article has received, and the thickness of each line reflects the number of co-citations received by two articles in other publications. The article by Schröder (2016) published in *Journal of Chemical Health & Safety* is included among the green nodes in Figure 5, having the highest number of links (504) and 32 co-citations. The article by Menard and Trant (2020) published in *Nature Chemistry* is included among the blue nodes with 486 links and 32 co-citations.

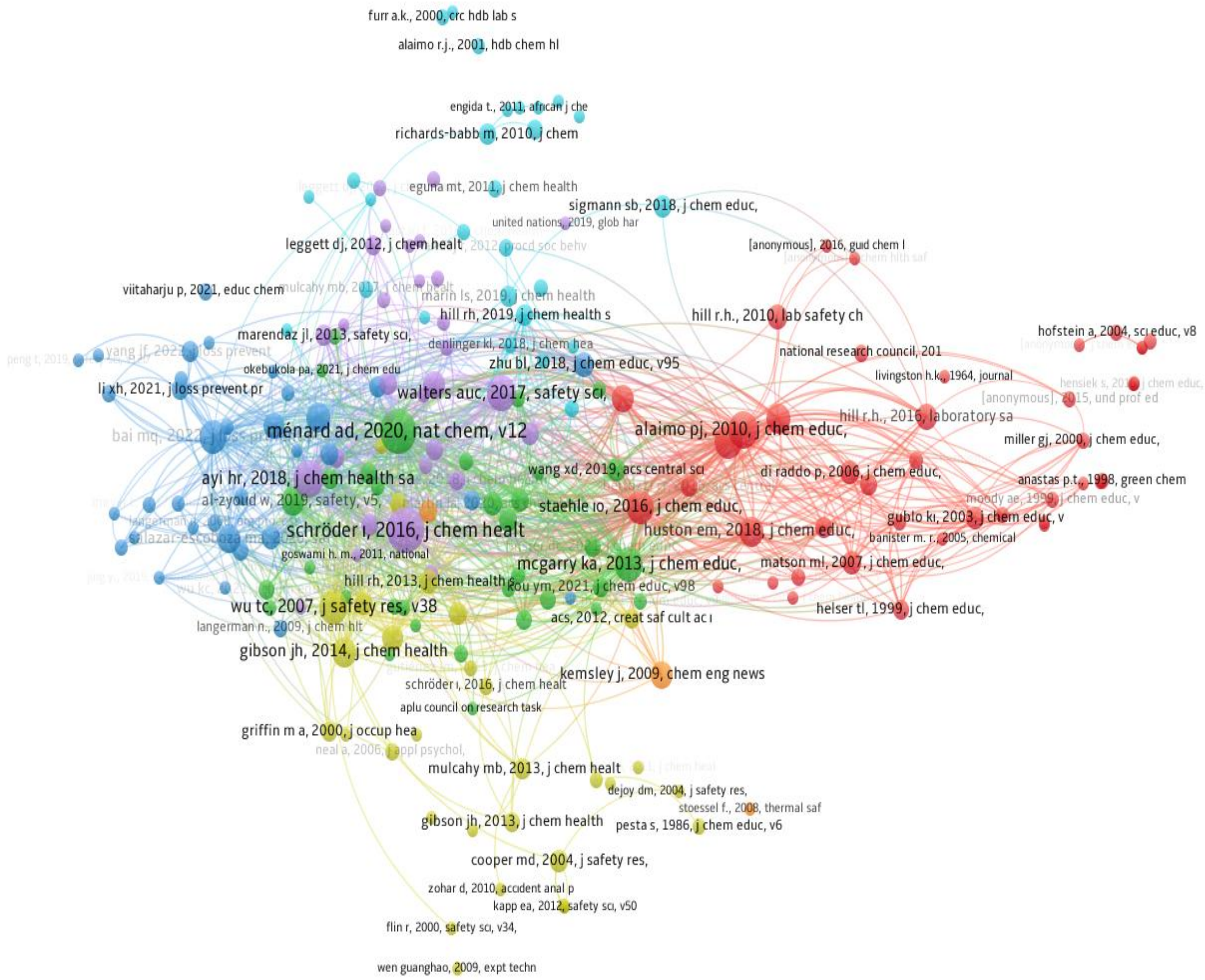



**Figure 5** Network Visualization Map of Co-citations

***Citations by countries***

Analysis was also conducted to evaluate the numbers of citations of different countries, and the results are shown in Figure 6 (Minimum number of publications of a country: 1; Minimum number of citations of a country: 1; of 60 countries, all 60 met these thresholds; Clusters: 14, Links: 159, Total link strength: 489). As seen in Figure 6, the countries with the highest numbers of citations are the United States with 1104 citations (99 publications), Canada with 286 citations (10 publications), and China with 244 citations (35 publications).

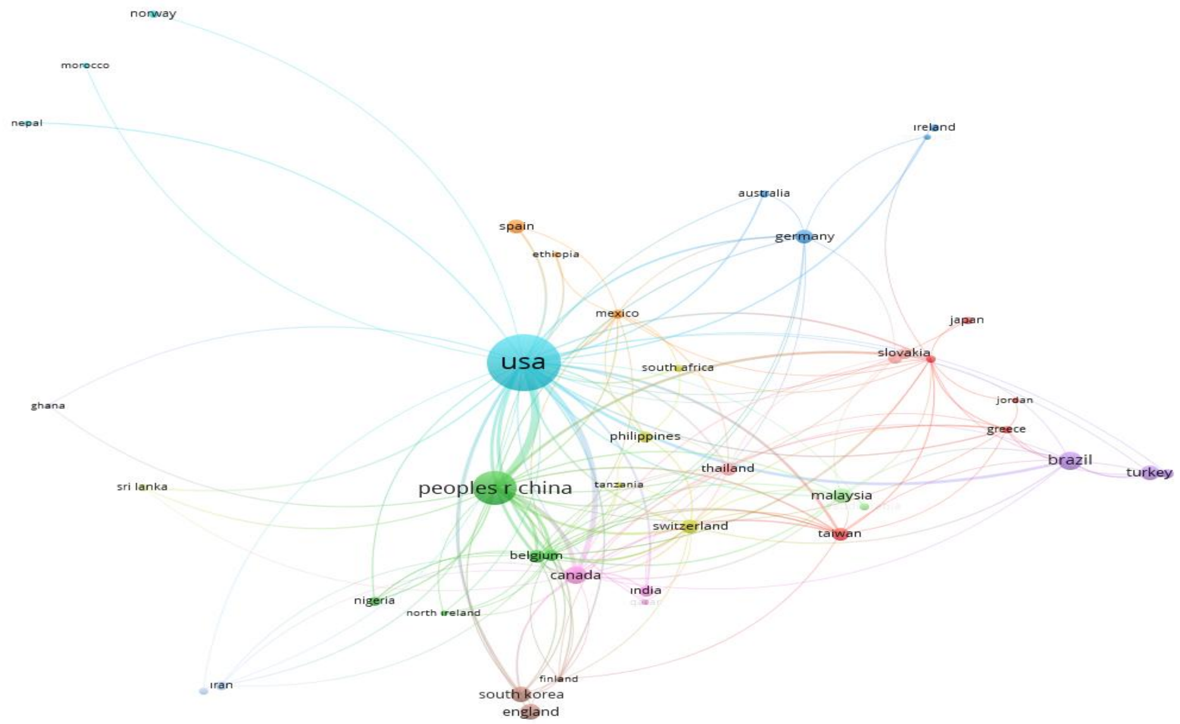



**Figure 6** Network Visualization Map of Citations by Countries

***Citations by institutions***

Analysis was conducted to evaluate the numbers of citations of different institutions (Minimum number of publications of an institution: 1; Minimum number of citations of an institution: 1; of 361 organizations, 285 met these thresholds; Clusters: 18; Links: 973; Total link strength: 1115). The colors in Figure 7 represent clusters of institutions and the lines represent the strength of cooperation among institutions. The most prolific institutions were found to be the University of California-Los Angeles with 180 citations (7 publications), Dow Chemical Co. with 118 citations (3 publications), the University of Minnesota with 110 citations (3 publications), Delft University of Technology with 109 citations (3 publications), and the University of Windsor with 108 citations (1 publication).

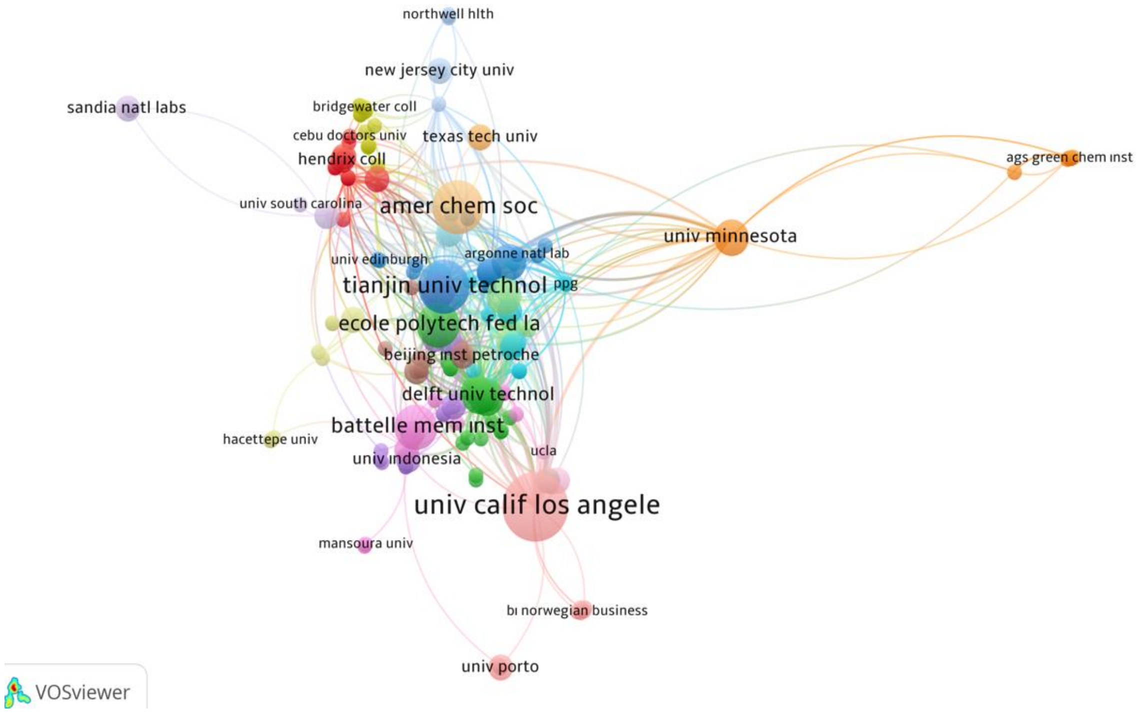



**Figure 7** Network Visualization Map of Institutional Cooperation

***Keyword analysis (co-occurrence of authors)***

The network structure obtained for the co-occurrence of authors according to keywords is presented in Figure 8 (Minimum number of occurrences of keywords: 1; of 572 keywords, all 572 met the threshold; Clusters: 33; Links: 1680; Total link strength: 2067). The most frequently used keywords were “laboratory safety” (37 occurrences), “safety hazards” (24 occurrences), “laboratory management” (23 occurrences), “safety” (18 occurrences), “laboratory instruction” (14 occurrences), and “risk assessment” (11 occurrences).

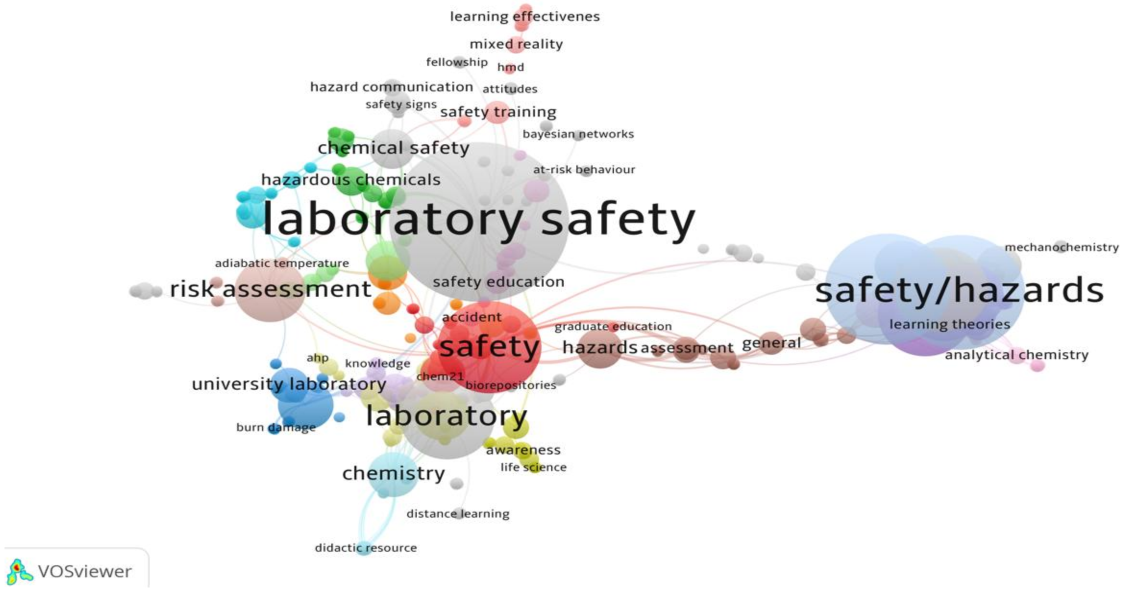



**Figure 8** Network Visualization Map of Keywords Co-Occurring Among Authors

## Conclusions and Suggestions

In this study, bibliometric analysis was conducted due to the multidisciplinary nature of laboratory safety and its significance in chemistry education. Safety is of the utmost importance in laboratory work. The findings obtained in this study were evaluated with a holistic approach. The results of the analysis are expected to guide future studies, contribute to the literature, and help other researchers.

This study examined 279 publications on the subject of laboratory safety. The documents reviewed were obtained using the WoS database and bibliometric analysis was performed. Considering the research on laboratory safety to date, the first publications appeared in 1965 and no increase in number was observed until 2009. In 2011, the number of annual publications increased and the highest numbers of studies were conducted between 2021 and 2023. The increase in the number of publications is an indicator of the importance of the field of laboratory safety at every level of education. Since the scope and applications of laboratory safety cover a wide spectrum, there are many different possible directions for advancements in this field. Yang et al. (2019) also stated that this area of research is quite young and open for improvement.

Reviewing the distribution of publications on laboratory safety according to authors, the most prolific researchers were found to be Nancy L. Wayne and Xiaoyan Wang, each with 6 articles, and Imke Schröder, James H. Gibson, and Xinglong Jin, each with 5 articles.

The institutions with the highest numbers of publications in the field of laboratory safety were found to be the University of California system and University of California-Los Angeles. Looking at the distribution of publications by institutions in Türkiye, the institutions with the highest numbers of publications in this field are Hacettepe University and Abant İzzet Baysal University.

The countries with the highest numbers of publications and the highest concentration of research were found to be the United States and China. These were followed by Brazil and Canada. On the other hand, in total, 7 relevant articles were published in Türkiye. It was thus found that the United States is considerably successful in terms of the number of publications and citations in this field, followed by China. Yang et al. (2019) emphasized that China's rise in this field of research should be taken into account.

It is of considerable value to identify publication and citation counts, and co-authorship relations among researchers studying laboratory safety to improve this area of study and help researchers acquire new information. When such information is available, we can identify the most productive researchers in laboratory safety. The contributions and the influence of specific researchers in a field constitute an essential indicator for the development of that field. The authors with the highest numbers of citations in the field of laboratory safety are James H. Gibson (160 citations), Nancy L. Wayne (143 citations), Imke Schröder (121 citations), and A.D. Menard and J.F. Trant (108 citations). When the network structure of the relationships among researchers who published co-authored articles in this field was examined, it was found that cooperation was concentrated around Nancy L. Wayne, James H. Gibson, and Imke Schröder.

The article published by Schröder (2016) in Journal of Chemical Health & Safety was cited 32 times, the article published by Menard and Trant (2020) in Nature Chemistry was cited 32 times, and the article published by Alaimo et al, 2010 in Journal of Chemical Education was cited 31 times. Additionally, these journals are among the most influential journals in the field.

Overall, the analysis conducted in this study showed that cooperation among authors in the field of laboratory safety is relatively weak. Yang et al. (2019) also reported that cooperation between countries, institutions, and authors is very low in this field.

Looking at the co-occurrence of keywords among the analyzed publications, the most frequently used keywords were “laboratory safety” (37 occurrences), “safety hazards” (24 occurrences), “laboratory management” (23 occurrences), “safety” (18 occurrences), “laboratory instruction” (14 occurrences), and “risk assessment” (11 occurrences).

This study examined laboratory safety in the context of the discipline of chemistry, and it was limited to studies indexed in the WoS database. Databases such as Google Scholar, ERIC, and SCOPUS were excluded from analysis. This study could be repeated using different databases. Conducting further bibliometric analyses of laboratory safety in science education could provide significant contributions to the field. In addition, the conduct and interpretation of bibliometric analysis studies by scientists who are experts in laboratory safety will make a great contribution to the field.

**Compliance with Ethical Standards**

*Disclosure of potential conflicts of interest*

No conflict of interest.

*Funding*

No payments have been received from any institution/person for this study.

*CRediT author statement*

The researchers took full responsibility for the research.

*Research involving human participants and/or animals*

This article did not involve human participants and/or animals.

## Laboratuvar Güvenliğinin Bibliyometrik Analizi: Kimya Disiplini için Önemi

**Özet:**

Laboratuvar güvenliği fen eğitimi ve özellikle kimya disiplini için oldukça önemli ve multidisipliner bir araştırma alanıdır. Bu çalışmada, kimya disiplininde laboratuvar güvenliğinin kapsamına ve uygulanmasına ilişkin gelişmeler, odak çalışma alanları ve eğilimler incelenmiştir. Araştırmada, betimsel yöntemlerden tarama modeli kullanılmış ve bibliyometrik analiz tekniği uygulanmıştır. Bu amaç kapsamında, 1965 ve 2024 yılları arasında laboratuvar güvenliği alanında Web of Science (WoS) Veri Tabanı'nda yayınlanan toplam 279 makale incelenmiştir. İncelemelerde 60 ülke, 279 araştırma, 51 yazar ve 43 konu kategorisi bulunmaktadır. Bulgular, VOSviewer ile bibliyometrik analiz yöntemi ile analiz edilmiştir. Elde edilen bulgulara göre ; araştırma alanı ile ilgili en fazla yayın yapılan ülkenin Amerika Birleşik Devletleri(ABD) olduğu, en fazla makale yayınlayan kurum University of California System , en fazla makale yayınlanan dergi Journal of Chemical Education, en fazla çalışma yapan araştırmacı Nancy L. Wayne ve Xiaoyan Wang olduğu ve bu araştırmacıları, Imke Schröeder, James H. Gibson ve Xinglong Jin izlediği tespit edilmiştir. Laboratuvar güvenliği konusunda çok sayıda bilim insanı araştırma yapmış olsa da yazarlar arasında işbirliğinin son derece zayıf olduğu saptanmıştır. En fazla kullanılan anahtar kelimeler laboratory safety, safety hazards laboratory management olduğu belirlenmiştir.

*Anahtar Kelimeler:* Bibliyometrik analiz, VOSviewer, laboratuvar güvenliği, kimya disiplini.

**References**

Agustian, H. Y., & Seery, M. K. (2017). Reasserting the role of pre-laboratory activities in chemistry education: a proposed framework for their design. *Chemistry Education Research,18*, 518-532. https://doi.org/10.1039/C7RP00140A

Akpullukcu, S. (2017). *The effects of laboratory safety professional development seminars implemented on science teachers: Laboratory safety knowledge levels* (Publication No. 486439) [Doctoral dissertation, Dokuz Eylül University]. Council of Higher Education Thesis Center.

Alaimo, P.J.,Langenhan, J.M., Joseph M.,Tanner, M.J., & Ferrenberg, S.M.(2010). Safety teams: An approach to engage students in laboratory safety, *Journal of Chemical Education*, *87*(8), 2825-2836. https://pubs.acs.org/journal/jceda8

Arslan, E. (2022). Bibliometric mapping in social sciences research using VOSviewer and an implementation, *Ankara Üniversitesi Sosyal Bilimler Dergisi*, *22*(2), 33-56. https://dergipark.org.tr/en/download/article-file/2864384

Atasoy, B. (2002). *Fen öğrenimi ve öğretimi.* Gündüz Eğitim ve Yayıncılık.

Ayana, U.C., Walters, W., Lawrence, N., & Jalsa K. (2017). Chemical laboratory safety awareness, attitudes and practices of tertiary students. *Safety Science, 96*, 161-171. https://doi.org/10.1016/j.ssci.2017.03.017

Ayi, H. R., & Hon, C. Y. (2018). Safety culture and safety compliance in academic laboratories: a Canadian perspective. *Journal of Chemical Health and Safety*, *25*, 6–12. https://doi.org/10.1016/j.jchas.2018.05.002

Block, J., & Fisch, C. (2020). Eight tips and questions for your bibliographic study in business and management research. *Management Review Quarterly*, (70), 307-312. https://doi.org/10.1007/s11301-020-00188-4

Çepni, S., Akdeniz, A. R., & Ayas, A. (1994). Fen bilimleri eğitiminde laboratuvarın yeri ve önemi (III): Ülkemizde laboratuvarın kullanımı ve bazı öneriler. *Çağdaş Eğitim Dergisi, 206*, 24-28.

De Bellis, N. (2009). *Bibliometrics and citation analysis: from the science citation index to cybermetrics*. Scarecrow. https://books.google.com.tr/books?id=ma4YjaKyM9cC&printsec=frontcover&hl=tr&source=gbs_ge_summary_r&cad=0#v=onepage&q&f=false.

Demir, E. (2016). *Investigating the physical conditions of the science laboratories and science teachers' level of knowledge about laboratory safety* (Publication No. 451833) [Master's thesis, Gazi University]. Council of Higher Education Thesis Center.

Gutiérrez-Salcedo, M., Martínez, M.Á., Moral-Muñoz, J.A., Herrera-Viedma, E., & Cobo, M. J. (2017). Some bibliometric procedures for analyzing and evaluating research fields. *Applied Intelligence*, *48*, 1275-1287. https://link.springer.com/article/10.1007/s10489-017-1105-y

Hill, R. H. (2007). The emergence of laboratory safety. *Journal of Chemical Health and Safety, 14*(3), 14–19. https://doi.org/10.1016/j.jchas.2006.10.001

Hofstein, A., & Lunetta, V. N. (2004). The laboratory in science education: Foundations for the twenty-first century. *Science Education*, *88*(1), 28-54. https://doi.org/10.1002/sce.10106

Hofstein, A., Navon, O., Kipnis, M., & Mamlok-Naaman, R. (2005). Developing students' ability to ask more and better questions resulting from inquiry-type chemistry laboratories. *Journal of Research in Science Teaching*, *42*(7), 791-806. https://doi.org/10.1002/tea.20072

Kurutkan, M. N., & Orhan, F. (2018). Bilim haritalama, bibliyometrik analiz ve kitap ile ilgili genel hususlar. M. N. Kurutkan & F. Orhan (Eds.), *Sağlık Politikası Konusunun Bilim Haritalama Teknikleri ile Analizi* (pp. 1-12). İksad. https://iksadyayinevi.com/wp-content/uploads/2020/02/SAĞLIK-POLİTİKASI-KONUSUNUN-BİLİM-HARİTALAMA-TEKNİKLERİ-İLE-ANALİZİ.pdf

McBurney, M. K., & Novak, P. L. (2002). What is bibliometrics and why should you care? *Proceedings of Professional Communication Conference, USA*, 108-114. https://doi.org/10.1109/IPCC.2002.1049094

Menard, A. D., & Trant, J. F. (2020). A review and critique of academic lab safety research. *Nature Chemistry*, *12*, 17-25. https://doi.org/10.1038/s41557-019-0375-x

Meyer, T. (2017). Towards the implementation of a safety education program in a teaching and research institution. *Education for Chemical Engineers*, *18*, 2–10. https://doi.org/10.1016/j.ece.2015.06.003

Miller G. J., Heideman S. A., & Greenbowe T. J. (2000). Introducing proper chemical hygiene and safety in the general chemistry curriculum. *Journal of Chemical Education*, *77*(9), 1185-1187. https://doi.org/10.1021/ed077p1185

Morgil, İ. F., & Yılmaz, A. (2000). Sicheres Arbeiten bei Lehramtsstudenten der Chemie in einer Türkischen Universitaet. *Chemkon*, *7*(4) ,198-203. https://doi.org/10.1002/ckon.20000070407

Nasrallah, I. M., El Kak, A. K., Ismil, L. A., Nasr R. R., & Bawab, W. T. (2022). Prevalence of accident occurrence among scientific laboratory workers of the public university in lebanon and the impact of safety measures. *Safety and Health at Work, 13*(2), 155-162. https://doi.org/10.1016/j.shaw.2022.02.001

Olewski, T., & Snakard, M. (2017). Challenges in applying process safety management at university laboratories. *Journal of Loss Prevention in the Process Industries*, *49*(B), 209−214. https://doi.org/10.1016/j.jlp.2017.06.013

Omidvari, M., Mansouri, N., & Nouri, J. (2015). A pattern of fire risk assessment and emergency management in educational center laboratories. *Safety Science, 73,* 34–42. https://doi.org/10.1016/j.ssci.2014.11.003

Pluess, D. N., Meyer, T., Masin, J., Mikulasek, P., & Ferjencik, M., (2016). Joint applicability test of software for laboratory assessment and risk analysis. *Journal of Loss Prevention in the Process Industries*, *40,* 234–240. https://doi.org/10.1016/j.jlp.2015.12.026

Schröder, I., Huang, D. Y. Q., Ellis, O., Gibson, J. H., & Wayne, N. L. (2016). Laboratory safety attitudes and practices: A comparison of academic, government, and industry researchers. *Journal of Chemical Health & Safety*, *23*(1), 12-23. https://doi.org/10.1016/j.jchas.2015.03.001

Seery, M. K., Hendra, Y., Agustian, H. Y., & Xinchi Zhang, X. (2018). A Framework for learning in the chemistry laboratory. *Israel Journal of Chemistry*, *59*, 546-553. https://doi.org/10.1002/ijch.201800093

Sigmann, S. (2018). Chemical safety education for the 21st century fostering safety information competency in chemists. *Journal of Chemical Health & Safety,25*(3), 17–29. https://doi.org/10.1016/j.jchas.2017.11.002

Stuart, R., & Toreki, R. (2014). Learning opportunities in three years of hazmat headlines. *Journal of Chemical Health* & *Safety*, *21*(2), 2−8. https://doi.org/10.1016/j.jchas.2013.11.002

Şener, F. (2018). *Scientific laboratory accidents reflected by printed and visual media that are occured in schools* (Publication No. 515967) [Master's thesis, Hacettepe University]. Council of Higher Education Thesis Center.

Tepe, M., & Tekbıyık A. (2019). Ortaokul fen bilimleri ders kitaplarının deney ve etkinlik güvenliği bakımından değerlendirilmesi, *Millî Eğitim, 48*(1), 223-240. https://dergipark.org.tr/tr/pub/milliegitim/issue/51765/673202

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

Walters, A. U. C., Lawrence, W., & Jalsa, N. K. (2017). Chemical laboratory safety awareness, attitudes and practices of tertiary students, *Safety Science, 96*, 161-171. https://doi.org/10.1016/j.ssci.2017.03.017

Wangdi, D., & Tshomo, S. (2016). Investigating chemical laboratory safety based on students' ability to recognize the common laboratory hazard symbols. *Educational Innovation and Practice. 1*(1), 41-52. https://www.researchgate.net/publication/325710061

Weil, M. (2016). The laboratory safety standard at 25: implementation of the standard through the chemical hygiene plan and the chemical hygiene officer – is it trickling down? *Journal of Chemical Health and Safety*, *23*(5), 31–40. https://doi.org/10.1016/j.jchas.2016.01.002

Yang, Y., Reniers, G., Chen, G., & Goerlandt, F. (2019). A bibliometric review of laboratory safety in universities. *Safety Science*, *120*, 14-24. https://doi.org/10.1016/j.ssci.2019.06.022

Yılmaz, A. (2023, September 6-8). *Laboratuvarda Güvenli Çalışma ve Öğrenme Ortamlarının Öneminin Ne Kadar Farkındayız?*. [Keynote], 8. Ulusal Kimya Eğitimi Kongresi, Marmara University, İstanbul, Türkiye.

Yılmaz, A. (2015). *Laboratuvarda Güvenli Çalışma*, Hacettepe Üniversitesi. https://library.hacettepe.edu.tr/hubooks/index.php?fn=read&key=429e6c

Yoon, D. (2021). Operational Process for R&D Laboratory Safety, *SAGE Open*, *11*(1),1-9. https://doi.org/10.1177/2158244020988853

Zhu, B., Feng, M., Lowe, H., Kesselman, J., Harrison, L., & Dempski, R.E. (2018). Increasing enthusiasm and enhancing learning for biochemistry-laboratory safety with an augmented- reality program. *Journal of Chemical Education*, *95*(10), 1747–1754. https://doi.org/10.1021/acs.jchemed.8b00116

Necatibey Faculty of Education Electronic Journal of Science and Mathematics Education
Vol. 18, Issue 2, December 2024, pp. 291-323. ISSN: 1307-6086

Research Article

# Mathematical Reasoning Activity: Compare, Generalize and Justify

**Tuba ÖZ[1], Zeynep ÇİFTCİ[2]**

[1] Atatürk University, Kazım Karabekir Education Faculty, Türkiye, tkaplan@atauni.edu.tr, http://orcid.org/0000-0003-0536-9360

[2] Atatürk University, Kazım Karabekir Education Faculty, Türkiye, zbayrakdar@atauni.edu.tr, http://orcid.org/0000-0002-3828-6230

Received : 28.06.2024 Accepted : 23.08.2024

Doi: https://doi.org/10.17522/balikesirnef.1506921

---

*Abstract* – The significance of mathematical reasoning skills is often highlighted in national and international curricula. In recent years, the process aspect of mathematical reasoning has been examined through comparison, generalization, and justification. Emphasizing these process abilities is crucial for creating learning settings that develop mathematical thinking and enhance teacher's understanding. This study assessed middle school students' comparation, generalization, and justification within reasoning activities. The participants were 27 sixth-grade students engaged in a mathematical reasoning workshop. The research data were gathered via a reasoning activity including three open-ended sub-problems addressed by the students. The data were analyzed using content analysis. The results showed that middle school students were capable of comparison, although they had difficulties in generalization and justification. Upon comprehensive evaluation, it was concluded that the number of students who completed these three steps cohesively was considerably low.

*Keywords*: Mathematical reasoning skill, comparing, generalizing, justification.

------------

Corresponding author: Zeynep ÇİFTCİ, zbayrakdar@atauni.edu.tr

## Introduction

Mathematical reasoning is essential for encouraging creative thinking and enhancing students' comprehension in the mathematics teaching process (Carpenter et al., 2003). Through the advancement of mathematical thinking, students will comprehend that mathematics is logical and comprehensible (Pengmanee, 2016). Mathematical reasoning necessitates procedural approaches to mathematical issues and the provision of explanations for answers (Waluyo et al., 2021). This ability is incorporated into the mathematics curriculum of several countries, interwoven with mathematics instruction at all grade levels (Brodie, 2010; Carpenter et al., 2003; Hunter, 2006; Kilpatrick et al., 2001; Visnovska & Cobb, 2009). Mathematical reasoning can be described as a universal mode of thought (Lithner, 2008). Mathematical reasoning entails selecting a strategy and implementing it to solve a problem (Säfström et al., 2024). Moreover, mathematical thinking is occasionally characterised as a general aptitude and a problem-solving instrument (Hjelte et al., 2020). Jeannotte and Kieran (2017) proposed a "concept blur" in the definition and conceptual framework of mathematical reasoning. In fact, mathematical reasoning is a concept that is sometimes difficult for teachers to recognize, let alone teach (Herbert & Williams, 2023). This uncertainty complicates the comprehension of how students might be supported in mathematical reasoning and the scientific research process associated with the idea (Hjelte et al., 2020). Jeannotte and Kieran (2017) defined mathematical reasoning as a communicative activity, either with others or internally, that facilitates the derivation of new mathematical statements from existing ones. They also pointed out that mathematical thinking has both structural and process aspects. Deductive, inductive, and abductive approaches characterize the structural dimension of mathematical reasoning, while the processes of identifying similarities and differences and verification methods represent the process dimension of mathematical reasoning (Jeannotte & Kieran, 2017). The structural component of mathematical reasoning includes more static situations, whereas the process component includes cognitive actions aimed at generating outcomes through inference. Consequently, the process dimension of mathematical thinking has garnered more attention in recent years (Widjaja et al., 2021; Geteregechi, 2020). Nevertheless, this process dimension of thinking is less examined in the literature (Jeannotte & Kieran, 2017). Numerous scholars have indicated that various mathematical reasoning processes, including comparing and contrasting, generalizing and justifying, conjecturing, persuading, and debating, are interconnected (Ellis, 2007; Jeannotte & Kieran, 2017; Lannin et al., 2011 as cited in Santos et al., 2022;

Stylianides, 2007). Jeannotte and Kieran (2017) characterized the process aspect of mathematical reasoning in two groups as "the process of investigating similarities and differences" and "the process of investigating justification ".They asserted that the process of investigating similarities and differences encompasses generalization, pattern recognition, hypothesis formulation, classification, and comparative analysis. They asserted that justification and evidentiary acts comprise the process of investigative verification, while exemplification serves as the action encompassing both processes.

Despite the inclusion of mathematical reasoning in the mathematics education standards of many countries, its application in classrooms remains seldom (Smit et al., 2023). Educators should comprehend their cognitive processes and inferential reasoning while addressing mathematical problems to gain insights into their students' learning (Yeşildere & Türnüklü, 2007). Consequently, evaluating students' mathematical reasoning processes is essential for comprehending learning and teaching contexts (Güler Baran, 2023). In this study, we will analyze the reasoning processes of "comparing," "generalizing," and "justification" by focusing on the process aspect of mathematical reasoning.

## Theoretical Background

### *Comparing*

The process dimension of reasoning involves the search for similarities and differences, which encompasses generalizing, making assumptions, identifying patterns, categorizing, and comparing (Widjaja et al., 2021). Comparing, as a reasoning process, is associated with the identification of similarities and differences within mathematical reasoning. This process involves the investigation of common and distinct features among mathematical objects, followed by the establishment of connections (Jeannotte & Kieran, 2017). Comparing is defined by Vale et al. (2017b) as the process of comparing and contrasting to identify a common aspect by recalling past information. These characteristics allow for comparisons to occur in various processes of mathematical reasoning (Jeannotte & Kieran, 2017). For example, consider the processes of pattern identification and conjecture formulation when investigating the similarities and differences among mathematical objects. Conjectures require the comparison of specific examples; likewise, identifying patterns necessitates the comparison of situations or examples (Pedomente, 2002). Pattern identification represents a progression beyond mere comparison. Comparison is confined to the construction of a narrative regarding similarities and differences (Jeannotte & Kieran, 2017).

Comparing is closely linked to generalizing during the examination of similarities and differences (Jeannotte & Kieran, 2017). Identifying patterns and similarities within the context during the comparison process serves as the foundation for generalization (Melhuish et al., 2020). The enquiries "What is the same?" and "What is different?" regarding a mathematical scenario necessitate a comparison and contrast of actions. The ability to distinguish critical aspects from non-critical ones is essential for addressing these enquiries and is crucial for facilitating generalization (Lo & Marton, 2012).

***Generalizing***

Generalizing is one of the indicators evaluated in the process dimension of mathematical reasoning skills (Jeannotte & Kieran, 2017). Generalizing is employed by mathematicians and mathematics educators to denote both a process and an outcome (Harel & Tall, 1991). A formal rule created from a generalizing job is termed a generalizing, which is a result (Chua, 2013); conversely, when a common or unifying quality is sought among a class of objects, it is seen as a process (Venenciano & Heck, 2016). Nevertheless, when a generalizing statement emerges from the generalizing process, both the method and the output may be regarded collectively (Yerushalmy, 1993, as cited in Oflaz, 2017). Ellis (2007) posits that generalization is a dynamic process wherein learners participate in at least one of the following activities: discovering commonalities across examples, broadening their thinking, or deriving overarching conclusions from particular instances. Numerous academics have characterized generalizing by highlighting its inferential and extensional dimensions. Jeannotte and Kieran (2017) describe generalizing as the process of deriving a link between a collection of mathematical entities or elements from a subset of that collection. Mason et al. (2010) define generalizing as the extension of outcomes derived from mathematical reasoning and problem-solving to a broader context. Lannin (2005) describe generalizing as the activity of contemplating analogous and ongoing occurrences within the broadest context. Kaput (1999) described generalizing as the identification of common characteristics among sample scenarios and the organization of communication and reasoning into a coherent pattern, structure, or connection (Kaput, 2008). The process of generalization is founded on recognizing patterns, delineating commonalities, and correlating analogous materials. The crucial aspect of this approach is not to identify the parallels between occurrences but to broaden and modify these similarities (Ayber, 2017). For instance, asserting that the elements in the sets 1, 3, 5, 7, and 9 are odd and increment by two exemplifies generalization (Hargreaves et al., 1998). The crucial aspect is to provide an explanation that transcends the dataset concerning the regularity of numerical features.

Mason's (1996) assertion that courses devoid of generalizations and assumptions do not constitute mathematics lessons, irrespective of their designation, underscores the significance of generalization in mathematical education. The capacity to generalize allows pupils to engage in systematic thinking and apply principles to specific scenarios (Venenciano & Heck, 2016). Moreover, generalization is a mathematical cognitive process that facilitates students' comprehension of symbolic representations and the establishment of connections with their existing arithmetic knowledge (Lannin, 2005). Generalizing involves utilizing mathematical meanings and relationships to construct accurate assumptions regarding mathematical structures (Melhuish et al., 2020). Generalization prompts the individual to address the enquiries: "What is probable (assumption), why is it valid (justification), and in what context is it applicable (general framework)?" Mason et al. (2010). Consequently, generalization is inseparable from the validation of the derived assertion (Lannin, 2005). While the formulation of a forecast, whether verbal or symbolic, serves as adequate evidence for generalization (Chua, 2013), it remains only a prediction until its correctness is substantiated (Watson, 1980). This research defines generalization as the extrapolation of observed similarities (relations or qualities) from a sample context to a broader context.

### *Justifying*

Justifying, substantiation, and formal proof fall under the category of verification (Widjaja et al., 2021). Verification pertains to critical functions including systematization, communication, integration, creation, and dissemination of new knowledge (Staples et al., 2012). Hanna (2000) identifies two primary functions of validation: to demonstrate truth and to elucidate the reasons for its truth. While it is often straightforward to ascertain "what," understanding "why" is considerably more complex. Addressing the why question necessitates a compelling justification (Mason et al., 2010). Mathematical justification necessitates an examination of existing knowledge and an assessment of the validity of assertions (Staples et al., 2012). Justification, which encompasses the arguments employed to validate and persuade, extends beyond mere explanation (Carpenter et al., 2003; Stebbing, 1952).

"Justifying" refers to the process of validating the truthfulness of information without engaging in a comprehensive proof process (Jeannotte & Kieran, 2017). While justifying serves the same purpose as proof and proving—assessing the truth of a statement—it diminishes the emphasis on the necessary level of formality and specificity that proof entails. This approach facilitates access to pertinent concepts while ensuring that no ideas are overlooked (Staples & Newton, 2016). A key distinction between mathematical proof and mathematical justification is that justifications do not require logical completeness (Jaffe, 1997). Justifying entails employing mathematics to convince oneself or others, irrespective of the completeness of the argument or its acceptance as incontrovertible evidence by the mathematical community (Lesseig, 2016). Melhuish et al. (2020) redirected the focus on student arguments from their completeness or correctness to the process of justification, encouraging greater student engagement with justification. Justification need not be formal or accurate; however, it remains a mathematical reasoning process (Lannin et al., 2011, as cited in Santos et al., 2022).

A strong mathematical justification should effectively address the question of "Why?". Addressing the why question elucidates the background of students' knowledge (Özmusul, 2018). The act of "justifying" enhances students' comprehension of mathematical concepts and aids them in uncovering the rationale behind mathematical principles, as well as substantively articulating their disagreements (Hanna, 2000). Justification serves as an effective learning practice and pedagogical instrument, enhancing students' comprehension of mathematics and facilitating mathematical processes (Staples et al., 2012). Justification allows students to comprehend mathematical concepts and to persuade others of the validity of the procedures, strategies, assumptions, or generalizations they employ (Carpenter et al., 2003; Dreyfus, 1999; Lannin et al., 2011 as cited in Santos et al., 2022; Lannin, 2005; Pedemonte, 2007).

This study defines justifying as the process of persuading the researcher by elucidating the validity of generalizations derived from observed commonalities (relationships or shared characteristics) within the sample context.

## The Objective and Significance of the Research

Mathematical reasoning skills are crucial for attaining mathematics learning objectives (Putra et al., 2020). The achievement of mathematics learning objectives has elevated the significance of mathematical literacy. A primary objective of the mathematics curriculum is to cultivate students capable of 'developing and successfully utilizing mathematical literacy abilities' (Ministry of National Education [MoNE], 2018). As stated in the OECD 2022 report, the Program for International Student Assessment (PISA) 2022 and Trends in International

Mathematics and Science Study (TIMSS) 2022 both investigated students' mathematical reasoning skills when they made the framework for the mathematical literacy assessment. The PISA (2022) Turkey Report indicates that Türkiye's reasoning performance is below the OECD average. Consequently, there is a must to enhance reasoning abilities in our country. National and international curricula emphasize the necessity of creating conducive circumstances for the cultivation of mathematical reasoning skills. Indicators essential for students to develop reasoning skills include 'defending the validity and truth of inferences,' 'formulating logical generalizations and inferences,' and 'articulating and applying mathematical patterns and relationships when analyzing a mathematical context' (MoNE, 2013).

The National Council of Teachers of Mathematics (NCTM) (2000) asserts that cultivating students' thinking relies on certain assumptions and principles and that students should be motivated to defend and formulate assumptions. Much research has indicated that justification and generalization are essential across all grade levels and are pivotal to the learning process (Blanton & Kaput 2003; Carraher et al., 2006; Ellis 2007; Lannin 2005) as quoted in Melhuish et al. (2020). By analyzing students' generalization of mathematical concepts, educators can discern the extent of their conceptual comprehension. For secondary school students, comprehending generalization is crucial for enhancing conceptual knowledge (Angraini, 2023). Justification is essential for students to comprehend significant mathematical structures, concepts, and procedures in the classroom (Thanheiser et al., 2021). Staples et al. (2012) underscored the significance of centering justification as a pedagogical activity and asserted that it should be incorporated into the K-12 curriculum. Furthermore, it is asserted that reasoning and proof should be included in all educational processes beginning from early life (Harel & Sowder, 2007; NCTM, 2000).

An examination of the literature reveals that studies are exploring mathematical reasoning skills from diverse perspectives (Bragg & Herbert, 2018; Çiftci, 2015; Çoban, 2010; Francisco & Maher, 2011; Herbert, 2014; Herbert & Bragg, 2021; Herbert & Williams, 2023; Herbert et al., 2022; Lannin, 2005; Loong et al., 2018; Marasabessy, 2021; Mata-Pereira & Ponte, 2017; Öz, 2017; Vale et al., 2017b). Jeannotte and Kieran (2017) noted that the process aspect of mathematical reasoning is inadequately addressed in the literature (Ellis, 2007; Herbert et al., 2022; Jeannotte & Kieran, 2017; Lannin et al., 2011; Lin & Tsai, 2016; Loong et al., 2018; Mason, 1982; Pedemonte, 2007; Peker, 2020; Stylianides, 2007, 2008; Widjaja & Vale, 2021). Emphasizing comparison, generalization, and justification activities, which signify reasoning skills, is crucial in designing learning environments to foster these skills and enhance teachers' knowledge. This study aimed to elucidate the current state of middle school students' comparison, generalization, and justification thinking processes during a mathematical task. Analyses of the students' reasoning and problem-solving approaches are useful in elucidating the nature of reasoning and the dynamics of the processes involved (Serrazina et al., 2024). Selecting suitable tasks or problems that will elucidate the thinking processes under examination is crucial. Vale et al. (2017a) asserted, "What else could it be?" activities such as "Which one does not belong?" can offer thinking possibilities across many mathematical ideas and different primary school levels (Small, 2011). The task “What else might it be?” was employed in our study to elucidate the thinking processes of middle school students transitioning from elementary school. The research is crucial for assessing the comparison, generalization, and justification processes of secondary school students in our country. The study's conclusions are significant since they offer insights for teachers and academics. The study aimed to address the subsequent research questions:

- How are the comparing processes of secondary school students?
- How are middle school students' generalizing processes?
- How are the justifying processes of middle school students?

## Method

### Research Design

As the study's objective is to describe the reasoning processes of sixth-grade students by implementing a reasoning activity that involves contrasting, generalizing, and justifying processes within the scope of the research, a case study, which is a qualitative research approach, was adopted. A case study is a research method favored for addressing "how" and "why" questions, particularly in contexts where researchers lack control over events or phenomena (Yin, 2009).

The study involved 27 sixth-grade students from two public schools who voluntarily engaged in a mathematical reasoning workshop organized by the researchers. The schools situated in the Eastern Anatolia Region of Türkiye exhibit similarities in socio-economic terms. The sample included 22 female students and 5 male students. The selection of 6th-grade students was based on the expectation that they would possess foundational skills related to numerical phenomena.

**Data Collection Tools**

The study utilized Small's (2011) "What else could it be?" activity. The Turkish version of the activity in the data collection tool received support from a language expert. Furthermore, the perspectives of three researchers with expertise in mathematics education were obtained regarding the implementation of the activity in the study. This activity enables students to compare, generalize, and justify, as the numerical set in the question stem encompasses multiple relationships and shared characteristics. The activity comprises three open-ended sub-questions. The initial inquiry of the activity is, "These numbers (30, 12, 18) belong together or not because…". The subsequent question is, "Other numbers that belong with this group are…". The third question is "How do you know that all these numbers belong and fit with your reason? Use words numbers or drawings to explain".

**Data Collection Process**

The data were collected during the mathematical reasoning workshop. The participants were informed in advance of the workshop's date, time, and location. The university provided transport support for students to attend the workshop. The workshop took place in a meeting room that accommodated students comfortably, facilitating ease of writing. Students engaged in the workshop alongside their teachers. The comfort of students in the environment is crucial for their active participation in the researchers' directives. The participants were required to provide written responses independently. The preference for written data collection was due to the presence of students from two different schools in the workshop, as it was anticipated that they might feel uncomfortable expressing themselves verbally. During the workshop, an overview was provided for the initial 15 minutes to enhance awareness of mathematical reasoning skills. Subsequently, the participants were allotted 20 minutes to engage in the "What else could it be" activity. The participants' responses were collected and retained for analysis after the session.

**Data Analysis**

Content analysis was implemented to analyze the research data. The data were analyzed through continuous comparison and grouping of relevant information. The researchers analyzed the data to generate codes, categories, and themes. The researchers collaboratively finalized the analysis process and achieved consensus on all code categories and themes. The researchers drew inspiration for naming certain codes, categories, and themes from the works of Ellis (2007), Vale et al. (2017b), and Widjaja et al. (2021). Two weeks post-analysis, the data underwent re-evaluation, leading to the finalization of the analysis.

## Role of the Researcher

Before the implementation process, researchers invited teachers and students to participate in the workshop voluntarily. At the outset of the workshop, the researchers provided an overview of mathematical reasoning skills. The nature of mathematical reasoning skills necessitates the frequent use of "why" and "if..." questions. Consequently, an effort was made to highlight how the "if..." structure underpins these reasoning skills. Subsequently, the researchers provided the data collection instrument to the students and remained available in the environment to address any potential enquiries from the students. In addressing potential student enquiries, efforts were made to avoid directing the students. The researchers conducted data arrangement, preparation for analysis, and the analysis process following data collection.

## Credibility, Transferability, Consistency, and Verifiability of the Research

The research's credibility was established through a comprehensive presentation of the methodologies employed, management of researcher biases (by engaging with an unfamiliar group), and data analysis conducted by two expert mathematics educators. The research stages were described in detail to enhance the transferability of the study. Using direct quotes and comparing the data to one another helped to verify consistency. The confirmability of the study was established through a detailed explanation of the analysis method, comprehensive descriptions of the participants, data collection tools, data storage procedures, and the researcher's role.

## Ethical Issues

The students, accompanied by their teachers, participated in the workshop, with verbal consent obtained for their involvement in the activities. In the data collection phase, participants were instructed to articulate their ideas without restriction. In the research, the actual names of the students who participated in the workshop were not utilized; instead, the student names were assigned codes such as S1 and S2. The data were transferred directly and unaltered.  Ethics committee approval was secured (Atatürk University Ethics Committee 05.07.2023/7).

## Findings

The analysis of data concerning the comparison, generalization, and justification processes—indicators of students' mathematical reasoning skills—was elucidated through tables, followed by examples that illustrate the diversity of analyses presented.

### Findings Related to Students' Comparing Process

The comparative situations of the students were derived from analyzing their responses to the question**,** "These numbers (30, 12, 18) belong together or not because**…”** in the data collection instrument. The responses to this question should demonstrate recognition of features such as magnitude, order, place value, multiples, factors, and the classification of numbers as odd or even. Table 1 displays the codes, categories, and themes obtained from the evaluation of the students' responses.

**Table 1** Codes, Categories, and Themes for the "Comparing" Process

| Themes | Categories | Codes | Participants |
|---|---|---|---|
| Recognizing the relationship | Addition/ subtraction relationship | Addition relationship | S6, S7, S8, S9, S10, S11, S15, S23, S24, S25 |
| | | Subtraction relationship | S1, S6, S8, S9, S17, S22, S23, S24 |
| | Pattern finding | Pattern finding | S7, S10, S13, S15, S19 |
| | | Inability to find a pattern | S14 |
| Recognizing the common features | Multiples/ factors | Having factors of 2,3,6 | S1, S2, S4, S8, S10, S13, S18, S20, S21, S22, S23, S24, S25 |
| | | Being a multiple of 6 | S7, S12, S15, S19, S26 |
| | | Having factors of 2,3 | S3, S7, S16, S27 |
| | | Having factors of 3,6 | S9 |
| | | Having a factor 2 | S5 |
| | Even/ odd number | Being an even number | S1, S2, S4, S8, S11, S12, S17, S18, S19, S20, S21, S22, S23, S24, S25, S26, S27 |
| | | The units digits are even and the tens digits are odd | S17 |
| | Digit value | Being two digits | S19, S20, S17 |
| | | Increasing the number of ones digits | S11 |

Analysis of Table 1 reveals that all students identified at least one relationship or common feature. The predominant relations identified were the "addition/subtraction" relation within the theme of recognizing relations and the "multiples/factors" common feature in the theme of recognizing common features. Analysis of the codes reveals that the group elements exhibit common characteristics: they are even, consist of multiples of 2, 3, and 6, and demonstrate an "addition/subtraction" relationship among them. Some students identified multiple relationships or common features, resulting in a total frequency that exceeded the number of participants. Figure 1 presents selected excerpts from the student responses.

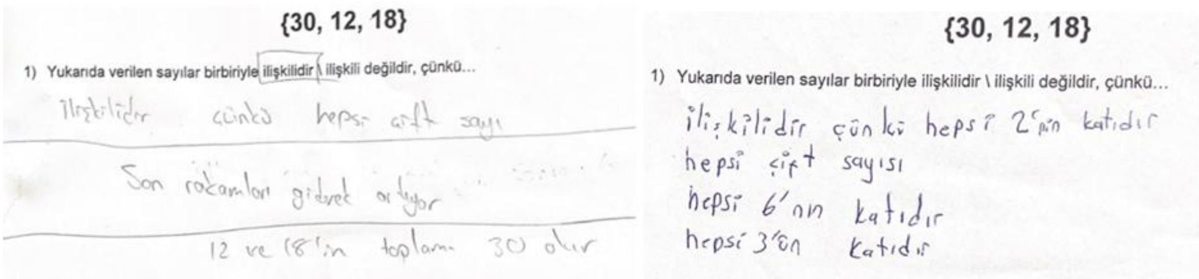



**Figure 1** Respectively S11 and S2's Answers About Comparing

Figure 1 illustrates S11's assertion regarding the additive relationship among the group elements, alongside S2's claim that the group elements possess factors of 2, 3, and 6. The relationship indicating that the sum of the numbers 12 and 18 in the group equals 30 was identified by numerous students through comparison. Furthermore, as indicated in S2's response, numerous students identified a shared characteristic through comparison, noting that all numbers within the group are multiples of 2, 3, and 6, and are classified as even numbers.

**Findings Related to Students' Generalizing Process**

The students' generalizing capacity was assessed through an analysis of their responses to the question, " Other numbers that belong with this group are…", within the data collection tool. The provided answers (30, 12, 18) indicate the necessity to extend the number set by incorporating additional values. All answers identifying the relationships or common features established by the student for the number group (30, 12, 18) were categorised under the theme "Expanding the group for the determined relationship/common feature."  Answers that failed to consider all relationships or common features identified by students in the generalisation step were categorised under the theme "Inability to expand the group for the determined relationship/common feature." Table 2 presents the codes, categories, and themes derived from the analysis of student responses.

**Table 2** Codes, Categories, and Themes for the "Generalizing" Process

| Themes | Categories | Codes | Participants |
|---|---|---|---|
| Expanding the group for the determined relationship/common feature | Expanding the group through a single relationship/common feature | Having a factor of 2 | S5 |
| | | Addition relationship | S6 |
| | Expanding the group over multiple relationships/common features | Being even and having a factor of 2 | S21 |
| | | Being even and having multipliers 2,3 | S27 |
| | | Being a multiple of 6 and being an even number | S26 |
| Inability to expand the group for the determined relationship/common feature | Forming a group that provides some of the determined relationships/common features | Having factors of 2,3,6 and being an even number | S1, S2, S8 |
| | | Addition relationship | S9 |
| | | Pattern finding | S10 |
| | | Being an even number | S17, S18, S20 |
| | | Even number and two-digits | S19 |
| | | Addition relation and having factors of 2, 3, 6 | S23, S24 |
| | Creating separate groups that fulfill some of the determined relationships/common features | Having a factor of 2 or 3 | S3, S4 |
| | | Being an even number | S4 |
| | Creating separate groups that fulfill each of the determined relationship/common features | Addition/subtraction relationship | S7, S11, S14, S15, S22, S25 |
| | | Being an even number | S11, S12, S22, S25 |
| | | Increasing the number of ones digits | S11 |
| | | Being a multiple of 6 | S7, S12, S15 |
| | | Having factors of 2, 3, 6 | S7, S13, S25 |
| | | Pattern finding | S7, S13, S14, S15 |
| | | Having a factor of 2, 3 or 6 | S16, S22 |

Analysis of Table 2 reveals that only five students expanded the group based on the relationships or common features identified during the comparison process. The other students were unable to identify the relationships or common features collectively and attempted to expand the group based on one or a few of these features. The students experienced challenges in expanding the number group by integrating all related features identified during the comparison process. However, they did not encounter difficulties in adding new numbers associated with a specific feature. During the analysis, it was observed that students who recognized both "addition relationship" and "pattern finding" relations, or the combination of "addition relationship," "pattern finding," and "being two digits" relations/common features simultaneously in the comparing phase, were unable to generalize all of these relations/common features. A student who extends the group based on the addition relation is unlikely to extend the group according to the -18+6 pattern rule. A student applying the -18+6 pattern rule to expand the group cannot achieve a superset that meets the criterion of being two-digit numbers. For this reason, students are expected to choose one of these properties and make a generalization accordingly. Figures 2 and 3 present excerpts from the students' responses.

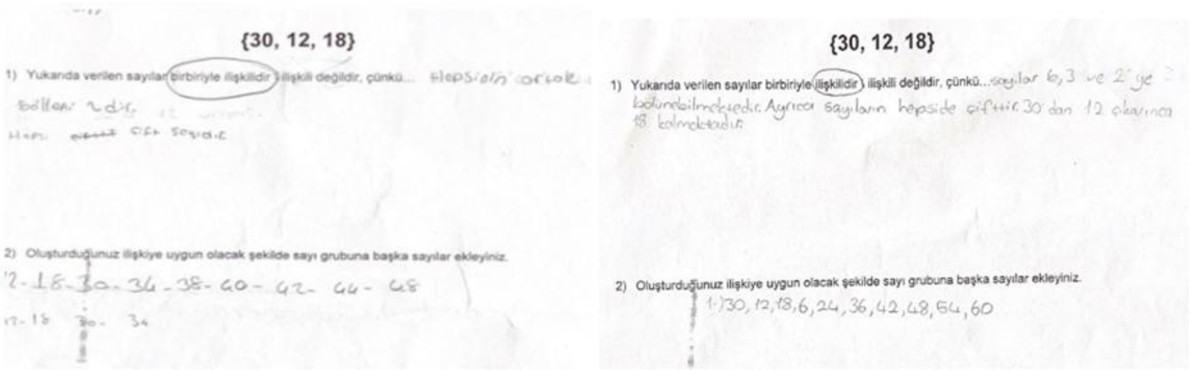



**Figure 2** Respectively S21 and S1's Answers About The Generalizing Process

Figure 2 illustrates that S21 successfully generalized the number group. S21 identified the shared characteristics of "being an even number" and "having a factor of 2," subsequently incorporating additional numbers into the group that exhibited both traits, thereby broadening the numerical set. S1, conversely, was unable to generalize the numerical group. S1 identified the shared characteristics of "having factors of 2, 3, 6" and "being an even number," as well as the "subtraction relationship" during comparisons. However, in the generalizing phase, S1 focused solely on the common features of "having factors of 2, 3, 6" and "being an even number," neglecting the "subtraction relationship." Consequently, the relations and common properties identified by S1 for the specified number group do not precisely align with those of the supergroup he attempted to establish, thus he cannot be regarded as having generalized the set.

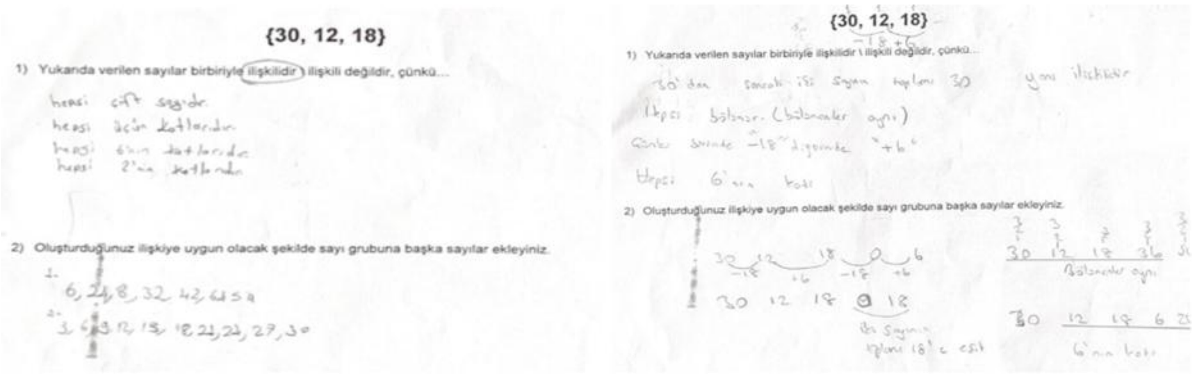



**Figure 3** Respectively S4 and S7's Answers About the Generalizing Process

Figure 3 illustrates that S4 was unable to generalize the number group. S4 identified the shared characteristics of "being an even number" and "having factors of 2, 3, and 6" during comparison, subsequently creating a distinct group by incorporating numbers that exhibited some of the common features he discerned while generalizing. Although S4 identified "being a multiple of 6" as a common feature during comparisons, he failed to incorporate this feature in his generalizations. S7 was unable to generalize the number group. S7 observed the connections between "pattern finding" and "addition relationship," as well as the shared characteristics of "having factors of 2, 3" and "being a multiple of 6." While making comparisons, S7 attempted to expand the group individually for each scenario, neglecting to consider all relationships and common features collectively. Upon careful analysis, it is evident that the "addition relationship" and "pattern finding" relations cannot be simultaneously established. Consequently, the student should either refrain from writing one of the relations or indicate that the two relations cannot be simultaneously achieved.

**Findings Related to Students' Justifying Process**

The students' capacity for justification was assessed through an analysis of their responses to the prompt, “How do you know that all these numbers belong and fit with your reason? Use words numbers or drawings to explain”. Students are required to justify the expansion process undertaken based on the relationships and common features identified. Table 3 presents the codes, categories, and themes derived from the analysis of student responses.

**Table 3** Codes, Categories, and Themes for the "Justification" Process

| Themes | Categories | Codes | Participants |
| --- | --- | --- | --- |
| Inability to justify | No answer/irrelevant response | No answer/irrelevant response | S2,S16 |
| | Writing the specified relationship/common feature verbatim | Writing the specified relationship/common feature verbatim | S1, S3, S4, S6, S7, S8, S9, S11, S12, S13, S14, S17, S18, S19, S20, S22, S23, S24, S25, S26 |
| To be able to justify | To be able to make partial justification for the determined relationship/common feature | Pattern finding | S10 |
| | | Being a multiple of 6 | S15 |
| | | Having factors of 2, 3 | S27 |
| | To be able to make a justification for the determined relationship/ common feature | Being an even number and having factors 2, 3, 6 | S21 |
| | | Having a factor of 2 | S5 |

Analysis of Table 3 reveals that the majority of students were unable to justify. The students believed they justified by reiterating the relationship and common features identified. This situation may stem from the students' insufficient experience in justification. Participants' responses that the specified relationship or common feature characteristic verbatim without any justification, as well as those that provided entirely irrelevant answers, were categorized under the theme of "Inability to justify." A subset of students conducted the verification process for a single relationship or common feature. Only two students extended the group by utilizing the initially identified relationship/common feature and successfully justified this expansion. Participants' responses that justified were categorized under the theme "To be able to justify." The number of individuals capable of justifying is notably low. Figures 4 and 5 present excerpts from student responses concerning the justification process.

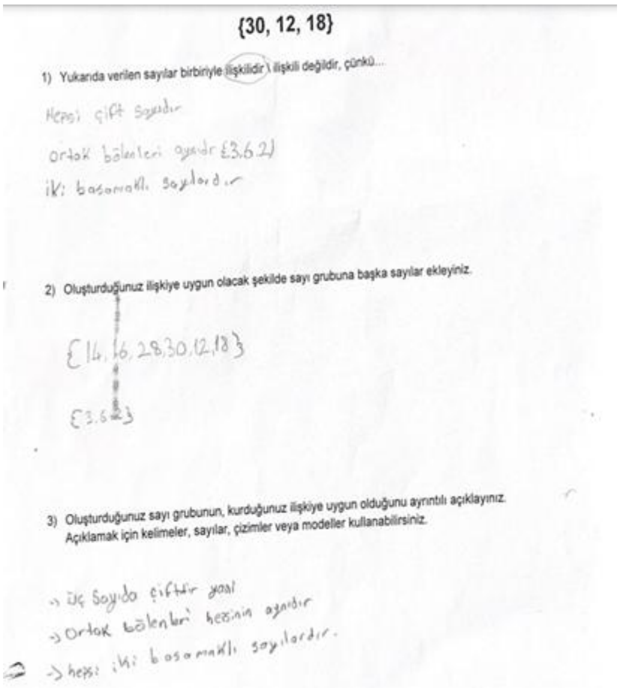



**Figure 4** S20's Answer about Justifying Process

As seen in  S20, there were many students who tended to write the determined relationship exactly. As can be seen in Figure 4, S20 formed a number group without using all of the relations/common features he identified in the comparing process. In the justification step, he wrote the relationship/common features he found without providing any justification.

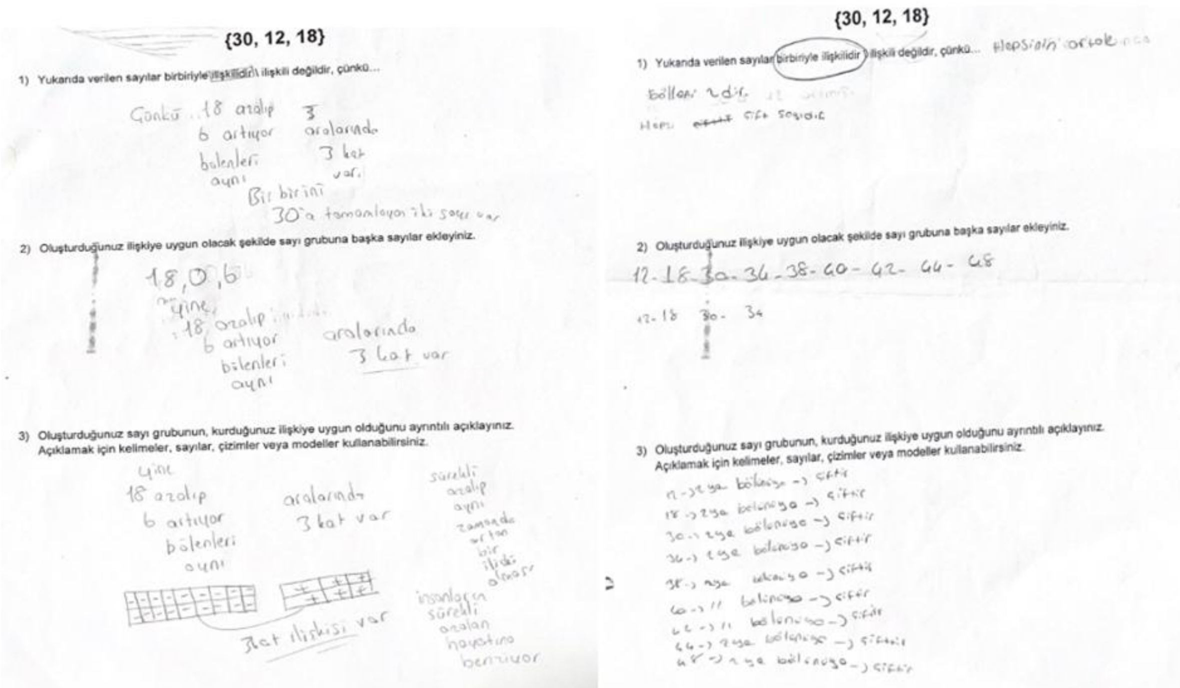



**Figure 5** Respectively S10 and S21's Answers about the Justifying Process

The answer of S10, which was evaluated in the category of "to be able to make partial justification for the determined relationship/common feature", is presented in Figure 5. As can be seen in the figure, S10 did not justify the relationship/common feature that he stated as "numbers complement each other by 30" among the relationship/common features he found, while he justified the other relationship/common features. The answer of S21 in the category of "to be able to make justification for the determined relationship / common feature" is given in Figure 5. S21 expanded the number group for the relationship/common feature he identified and was able to justify all the relationship/common features he found.

## Discussion

Our findings are essential for enhancing students' higher-level reasoning by identifying the common characteristics of this number group, extending the group to a broader set, and providing justification for their assumptions when faced with a number group. Upon examination of the students' responses, it was observed that they demonstrated at least one reasoning action in the processes of comparison, generalization, and justification. Upon comprehensive consideration, it was determined that the number of students who completed these three processes in a connected manner was relatively low. The research allowed students to articulate their thoughts freely through open-ended questions, mitigating grade-related anxiety. Vale et al. (2017a) underscore the fact that open-ended tasks that necessitate making assumptions about a shared characteristic offer an opportunity for reasoning actions.

The students demonstrated the ability to make comparisons to identify the relationships and common features among the numbers in the specified group. The students identified multiple relationships and common features within the given number group through their comparisons. The predominant relationship identified within the theme of "recognizing the relationship" is the "addition/subtraction" relationship. The investigation of the number group (30, 12, 18) primarily concentrated on the operations of addition and subtraction concerning establishing a relationship among the numbers. Additionally, some students sought patterns in their attempts to establish relationships between numbers. It is important to note that the number of students (7) attempting to identify a pattern was limited. The recognition of the "addition/subtraction" relationship by the majority (18) may be attributed to its minimal requirement for advanced reasoning in identifying common features among numbers. Jeannotte and Kieran (2017) asserted that identifying a pattern extends beyond mere comparison, as the process primarily highlights only similarities and differences. As an example of higher-level reasoning in the process of comparing and contrasting, Ellis (2007) and Vale et al. (2017b) also mentioned recognizing relationships that call for deeper thought rather than identifying relationships that are readily recognized. Our study indicates that students demonstrated a greater awareness of the addition/subtraction relation compared to the pattern-finding relation. Within the theme of "recognizing the common feature," it was noted that students identified the common features of "multiples/multipliers" and "even/odd numbers" more frequently than other common features. Students identified that the common divisors of the numbers in the group (30, 12, 18) are 2, 3, and 6. Although a minority of students identified only one or two common divisors, the majority successfully recognized all three common divisors. Furthermore, a notable characteristic that was often highlighted was the property of being an even number. Students recognized that every number in the group is an even number. Alongside these common properties, another category that some students emphasized is "digit value." According to Vale et al. (2017a), when students compare, they employ numerical information such as factors, multiples, place value, counting patterns, and number order. The potential to identify a relationship or common feature may be associated with students' prior knowledge of numerical phenomena. Consequently, students' capacity for comparison is linked to gaps in their prior knowledge.

The majority of students were unable to generalize the provided number group. It was observed that, during comparisons, the group could not be expanded by incorporating additional numbers related to the identified relationships or common features. Similarly, Rodrigues et al. (2021) highlighted that instructors and pre-service teachers struggle with scenarios requiring the generalization process, and Ersoy et al. (2017) also highlighted that secondary school students struggle with it. To achieve comprehensive generalization, it is essential to consider all identified relationships and common features holistically. Generalization necessitates that students recognize similarities and fundamental principles across diverse examples or contexts (Angraini, 2023). Malara (2012) describes the generalization process as a sequential cognitive activity that entails the analysis of specific instances and shared characteristics, subsequently applying these insights to all established common features. The groups formed by students who cannot generalize do not qualify as a superset when considering the established relationships or common features. The students demonstrated the ability to generalize across one or more relations or common features. It remains unclear whether students generalized over a single relation or common feature due to an inability to identify additional relations or because it was more straightforward to generalize in this manner. Students may encounter difficulties in the process of generalization, even when they recognize relationships or common features during comparison. This may lead them to concentrate on familiar relationships or features that they can easily manipulate.

The analysis of the justifying process revealed that students could not typically provide justifications. Only two students successfully generalized and justified the relationship or common feature identified within the number group during the comparison process. The findings align with previous research indicating a low incidence of students providing complete and convincing justifications (Özmusul & Bindak, 2022) and a general unfamiliarity among students with the practice of justifying their solutions (Reyes-Hernandez & Mooney, 2021). Furthermore, many studies indicate that students struggle with generalization and justification (Chazan, 1993; English & Warren, 1995; Knuth et al., 2002). When the answers of the students who could not justify are analyzed, it is noteworthy that the students tended to write the same relationship/common features they found. Students experienced challenges in generalizing the relationship or common feature identified within the number group to a broader context. Additionally, they struggled to establish and interpret the causal connections between the conjectures they developed. Understanding the rationale and methodology behind a task is crucial for the development of mathematical thinking in students, rather than merely executing tasks mechanically (Dikkartın Övez & İnce, 2024). Lins (2001) asserts that justifications offer insight into the overarching concept of generalization and its characteristics. Generalizing and justifying are closely related concepts (Ellis 2007; Kirwan 2015; Lannin 2005). Furthermore, justifying serves as a mechanism that aids students in uncovering the applications of various elements in mathematics, thereby enhancing their comprehension of mathematical concepts (Hanna, 2000). Students' inability to justify may be linked to their exposure to justifying activities within the learning environment and their underlying conceptual understanding. Assigning students tasks that facilitate the establishment of mathematical relationships, encourage discussion of their reasoning, and require justification within the learning environment allows for opportunities to justify and generalize (Staples & Newton, 2016; Stein et al., 2008). Bozkurt et al. (2017) indicated that secondary school teachers predominantly employed questions that elicited short answers in the classroom, while questions necessitating long answers and deeper comprehension, such as those involving justification and criticism/interpretation, were utilized to a lesser degree. Furthermore, a lack of emphasis on the rationale behind any procedure or phenomenon presented to students will diminish their engagement in hypothesizing, justifying, and generalizing activities (Mukuka et al., 2023). Jackson and Stenger (2024) emphasized that generalization should be conveyed through lessons designed to communicate overarching statements regarding the subjects studied in the classroom environment. It is essential for teachers to priorities comparing, generalizing, and justifying within their classrooms and to incorporate these elements into their teaching methodologies. Teachers are essential in facilitating students' engagement in mathematical reasoning (Ellis et al., 2019). Widjaja et al. (2021) highlighted the necessity of offering primary school students opportunities to engage in reasoning processes that involve comparing and contrasting, verifying conjectures, and generalizing. This study examined the process aspect of mathematical reasoning and concluded that the majority of middle school students were unable to complete the comparing, generalizing, and justifying processes holistically. When these processes were analyzed individually, the majority of students demonstrated the ability to make comparisons based on more superficial reasoning compared to other processes. Nonetheless, the proportion of students who engaged in the generalizing and justification processes necessitating higher-level reasoning significantly declined. However, what is expected here, is necessary to generalize the common features or relationships identified, to broaden the dataset, and to provide justification for this scenario.

The study's findings and results led to recommendations for researchers, educators, and program developers.

The process dimension of mathematical reasoning skills is gaining significance. Future studies should incorporate a broader range of research participants and methodologies. Given the extended duration necessary for students to develop a skill, further longitudinal studies are essential. Given the variability in students' modes of expression, it is beneficial to explore alternative assessment methods beyond solely evaluating their reasoning through written responses.

A conducive classroom climate is essential for fostering the development of mathematical reasoning skills in educational settings. Creating a supportive and non-judgmental environment is essential for facilitating student expression of thoughts. To address deficiencies in students' mathematical reasoning skills, it is essential to provide additional activities that facilitate reasoning opportunities and allow for individual reasoning development.

The acquisition of the mathematical reasoning skills, often implicitly integrated into curricula, can be emphasized more prominently. Teachers' professional development activities, as implementers of the curriculum, can be enhanced through more effective methods and instructional strategies for teaching this skill to students, facilitated by both in-service training and curricular improvements.

## Compliance with Ethical Standards

*Disclosure of potential conflicts of interest*

The authors have no relevant financial or non-financial interests to disclose.

*Funding*

The authors did not receive support from any organization for the submitted work.

*Credit author statement*

All authors contributed to the study's conception and design. Material preparation, data collection and analysis were performed by [Öz, T.] and [Çiftçi, Z.]. The first draft of the manuscript was written by [Öz, T.] and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

*Research involving Human Participants and/or Animals*

The participants were verbally informed.

Study-specific approval by the appropriate ethics committee for research involving humans and/or animals (The research study that underpins this publication was provided by Ataturk University, Registration number 05.07.2023/7).

## Matematiksel Akıl Yürütme Etkinliği: Karşılaştır, Genelle, Gerekçele

**Özet:**

Matematiksel akıl yürütme becerisinin önemi hem ulusal hem de uluslararası müfredatlarda sıklıkla vurgulamaktadır. Özellikle son yıllarda, matematiksel akıl yürütmenin süreç yönü, karşılaştırma, genelleme ve gerekçelendirme açısından ele alınmıştır. Bu süreç becerilerine odaklanmak, matematiksel akıl yürütme becerilerinin geliştirilmesi için öğrenme ortamlarının hazırlanmasında ve öğretmenlerin farkındalığının artırılmasında oldukça önemlidir. Bu çalışma ile ortaokul öğrencilerinin akıl yürütme etkinliği kapsamında karşılaştırma, genelleme ve gerekçelendirme durumları incelenmiştir. Araştırmanın katılımcılarını matematiksel akıl yürütme atölyesine katılan 6. sınıf seviyesindeki 27 öğrenci oluşturmaktadır. Araştırma verileri öğrencilerin cevaplandırdığı 3 açık uçlu alt problemden oluşan akıl yürütme etkinliği ile toplanmıştır. Veriler içerik analizi ile analiz edilmiştir. Araştırmanın sonuçlarına göre, ortaokul öğrencilerinin karşılaştırma yapabildiklerini fakat genelleme ve gerekçelendirme basamaklarında problem yaşadıklarını göstermektedir. Bir bütün olarak düşünüldüğünde ise bu üç süreci de bağlantılı bir şekilde tamamlayan öğrenci sayısının çok az olduğu sonucuna varılmıştır.

*Anahtar kelimeler*: Matematiksel akıl yürütme becerisi, karşılaştırma, genelleme, gerekçelendirme.

## References


Angraini, L. M., Larsari, V. N., Muhammad, I., & Kania, N. (2023). Generalizations and analogical reasoning of junior high school viewed from Bruner's learning theory. *Infinity Journal, 12*(2), 291-306. https://doi.org/10.22460/infinity.v12i2.p291-306

Ayber, G. (2017). *An analysis of secondary school mathematics textbooks from the perspective of fostering algebraic thinking through generalization* (Publication No. 463446) [Master’s thesis, Anadolu University]. Council of Higher Education Thesis Center.

Blanton, M., & Kaput, J. (2003). Developing elementary teachers’ algebra “eyes and ears”. *Teaching Children Mathematics, 10*(2), 70–77. https://doi.org/10.5951/TCM.10.2.0070

Bozkurt, A., Kılıç, P., & Özmantar, M. F. (2017). An investigation of the question types in mathematics instruction of middle school classrooms. *Yıldız Journal of Educational Research, 2* (1), 1-29. https://dergipark.org.tr/en/download/article-file/2088927

Bragg, L. A., & Herbert, S. (2018). What can be learned from teachers assessing mathematical reasoning: A case study. *Making waves, opening spaces: Proceedings of the 41st annual conference of the Mathematics Education Research Group of Australasia,* 178–185. https://files.eric.ed.gov/fulltext/ED592480.pdf

Brodie, K. (2010). *Teaching mathematical reasoning in secondary school classrooms*. Springer Science+Business Media.

Chazan, D. (1993). High school geometry students' justification for their views of empirical evidence and mathematical proof. *Educational Studies in Mathematics, 24*(4), 359–387. https://doi.org/10.1007/BF01273371

Çoban, H. (2010). *Investigating the relationship between the level of students? Using mathematical reasoning skills and using metacognitive learning strategies* (Publication No. 258052) [Master’s thesis, Gaziosmanpaşa University]. Council of Higher Education Thesis Center.

Carpenter, T. P., Franke, M., & Levi, L. (2003). *Thinking mathematically: Integrating arithmetic and algebra in elementary school.* Heineman.

Carraher, D. W., Schliemann, A. D., Brizuela, B. M., & Earnest, D. (2006). Arithmetic and algebra in early mathematics education. *Journal for Research in Mathematics Education*, *37*(2), 87–115. https://doi.org/10.2307/30034843

Chua, B. L. (2013). *Pattern generalisation in secondary school mathematics: students' strategies, justifications and beliefs and the influence of task features* [Doctoral dissertation, London University]. UCL Discovery.

Çiftci, Z. (2015). *The analysis of mathematical reasoning skills of middle school math teacher candidates* (Publication No. 418254) [Doctoral dissertation, Atatürk University]. Council of Higher Education Thesis Center.

Dikkartın-Övez, F. T., & İnce, İ. (2024). Examination of seven grade students'; pattern generalization processes and preferred, *International Journal of Education Technology and Scientific Researches*, *9*(26), 84-129. http://dx.doi.org/10.35826/ijetsar.728

Dreyfus, T. (1999). Why Johnny can't prove. *Educational Studies in Mathematics*, *38*(1), 85-109. https://doi.org/10.1023/A:1003660018579

Ellis, A. B. (2007). Connections between generalizing and justifying: Students' reasoning with linear relationships. *Journal for Research in Mathematics Education*, *38*(3), 194 - 229. https://doi.org/10.2307/30034866

Ellis, A., Özgür, Z., & Reiten, L. (2019). The teacher moves to support student reasoning. *Mathematics Education Research Journal*, *31*, 107-132. https://doi.org/10.1007/s13394-018-0246-6

English, L., & Warren, E. (1995). General reasoning processes and elementary algebraic understanding: Implications for instruction. *Focus on Learning Problems in Mathematics*, *17*(4), 1–19. https://eric.ed.gov/?id=EJ526526

Ersoy, E., Yıldız, İ., & Süleymanoğlu, E. (2017). A study upon the reasoning abilities of 5th graders. *Electronic Turkish Studies*, *12*(17), 179-194. http://doi.org/10.7827/TurkishStudies.11702

Francisco, J. M., & Maher, C. A. (2011). Teachers attending to students' mathematical reasoning: Lessons from an after-school research program. *Journal of Mathematics Teacher Education*, *14*, 49-66. https://doi.org/10.1007/s10857-010-9144-x

Geteregechi, J. M. (2020). *Investigating middle school students' mathematical reasoning in a connected in-school and out-of-school context* (Publication No. 28091313) [Doctoral dissertation, Syracuse University]. ProQuest Dissertations & Theses.

Güler Baran, H. (2023). *Investigation of mathematical reasoning processes of middle school students* (Publication No. 779335) [Master's thesis, Gazi University]. Council of Higher Education Thesis Center.

Hanna, G. (2000). Proof, explanation and exploration: An overview. *Educational Studies in Mathematics*, *44*(1), 5-23. https://doi.org/10.1023/A:1012737223465

Harel, G., & Tall, D. (1991). The general, the abstract, and the generic in advanced mathematics. *For the Learning of Mathematics*, *11*(1), 38-42. https://www.jstor.org/stable/40248005

Harel, G., & Sowder, L. (2007). Toward comprehensive perspectives on the learning and teaching of proof. In Lester, F. (Ed.), *Second handbook of research on mathematics teaching and learning*, *2*, 805-842.

Hargreaves, M., Shorrocks-Taylor, D., & Threlfall, J. (1998). Children's strategies with number patterns. *Educational Studies*, *24*(3), 315-331. https://doi.org/10.1080/0305569980240305

Herbert, S. (2014). *A framework for teachers' knowledge of mathematical reasoning.* Proceedings of the 36th annual conference of the Mathematics Education Research Group of Australasia, 702–705. https://www.researchgate.net/publication/281245963_A_Framework_for_Teachers'_Knowledge_of_Mathematical_Reasoning

Herbert, S., & Bragg, L. A. (2021). Elementary teachers' planning for mathematical reasoning through peer learning teams. *International Journal for Mathematics Teaching and Learning*, *22*(1), 24-43. https://doi.org/10.4256/ijmtl.v22i1.291

Herbert, S., Vale, C., White, P., & Bragg, L. A. (2022). Engagement with a formative assessment rubric: A case of mathematical reasoning. *International Journal of Educational Research*, *111*, 101899. https://doi.org/10.1016/j.ijer.2021.101899

Herbert, S., & Williams, G. (2023). Eliciting mathematical reasoning during early primary problem-solving. *Mathematics Education Research Journal*, *35*, 1-27. https://doi.org/10.1007/s13394-021-00376-9

Hjelte, A., Schindler, M., & Nilsson, P. (2020). Kinds of mathematical reasoning addressed in empirical research in mathematics education: A systematic review. *Education Sciences*, *10*(10), 289. https://doi.org/10.3390/educsci10100289

Hunter, J. (2006). Students in a mathematical community of inquiry: What do they think? In P. Grootenboer, R. Zevenbergen & M. Chinnappan (Eds.), *Proceedings of the 29th annual conference of the mathematics education research group of Australasia* (pp. 301–308).

Jackson, J.L., & Stenger, C.L. (2024). Methods of explicitly teaching generalization in the mathematics classroom and indicators of success: A systematic review. *International Journal of Education in Mathematics, Science, and Technology (IJEMST), 12*(4), 1109-1126. https://doi.org/10.46328/ijemst.4171

Jaffe, A. (1997). Proof and the evolution of mathematics. *Synthese*, *111*(2), 133–146. https://www.jstor.org/stable/20117623

Jeannotte, D., & Kieran, C. (2017). A conceptual model of mathematical reasoning for school mathematics. *Educational Studies in Mathematics*, *96*, 1-16. https://doi.org/10.1007/s10649-017-9761-8

Kaput, J. (1999). Teaching and Learning a New Algebra. In E. Fennema and T. Romberg (Eds.), *Mathematics Classrooms that Promote Understanding* (pp. 133–155). Routledge.

Kaput, J. J. (2008). What is algebra? What is algebraic reasoning? In Kaput, J. J., Carraher, D. W., & Blanton, M. L. (Eds.), *Algebra in the Early Grades* (pp. 5– 17). Routledge.

Kilpatrick, J., Swafford, J., & Findell, B. (2001). *Adding it up: helping children learn mathematics*. National Academy.

Kirwan, J. V. (2015). *Preservice secondary mathematics teachers' knowledge of generalization and justification on geometric numerical patterning tasks* (Publication No. 3722753) [Doctoral dissertation, Illinois State University]. ProQuest Dissertations & Theses.

Knuth, E. J., Slaughter, M., Choppin, J., & Sutherland, J. (2002). Mapping the conceptual terrain of middle school students' competencies in justifying and proving. In S. Mewborn, P. Sztajn, D.Y. White, H.G. Wiegel, R.L., Bryant,and K. Nooney (Eds.), *Proceedings of the 24th Meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education*. (Vol. 4, pp. 1693-1700).

Lannin, J. K. (2005). Generalization and justification: The challenge of introducing algebraic reasoning through patterning activities. *Mathematical Thinking and Learning*, *7*(3), 231-258. https://doi.org/10.1207/s15327833mtl0703_3

Lesseig, K. (2016). Conjecturing, generalizing and justifying: Building theory around teacher knowledge of proving. *International Journal for Mathematics Teaching and Learning*, *17*(3). https://doi.org/10.4256/ijmtl.v17i3.27

Lithner, J. (2008). A research framework for creative and imitative reasoning. *Educational Studies in Mathematics*, *67*(3), 255–276. https://doi.org/10.1007/s10649-007-9104-2

Lins, R. C. (2001). The production of meaning for algebra: A perspective based on a theoretical model of semantic fields. In R. Lins (Ed.), *Perspectives on school algebra* (pp. 37–60). Kluwer.

Lin, P. J., & Tsai, W. H. (2016). Enhancing students' mathematical conjecturing and justification in third-grade classrooms: The sum of even/odd numbers. *Journal of Mathematics Education*, *9*(1), 1-15. https://www.educationforatoz.com/images/Pi-Jen_Lin_2016.pdf

Lo, M. P., & Marton, F. (2012). Towards a science of the art of teaching: using variation theory as a guiding principle of pedagogical design. *International Journal for Lesson and Learning Studies*, *1*(1), 7–22. https://doi.org/10.1108/20468251211179678

Loong, E., Vale, C., Widjaja, W., Herbert, E.S., Bragg, L., & Davidson, A. (2018). Developing a rubric for assessing mathematical reasoning: A design-based research study in primary classrooms. Making waves, opening spaces: *Proceedings of the 41st annual conference of the Mathematics Education Research Group of Australasia*, 503-510.
https://www.researchgate.net/publication/334950414_Developing_a_Rubric_for_Assessing_Mathematical_Reasoning_A_Design
Based_Research_Study_in_Primary_Classrooms

Malara, N. A. (2012). Generalization processes in the teaching/learning of algebra: students behaviours and teacher role. In: Maj-Tatsis B., Tatsis K. (Eds.) *Generalization in mathematics at all educational levels*. 57-90.

Marasabessy, R. (2021). Study of mathematical reasoning ability for mathematics learning in schools: A literature review. *Indonesian Journal of Teaching in Science, 1*(2), 79-90. https://doi.org/10.17509/ijotis.v1i2.37950

Mason, J. (1982). *Thinking mathematically*. Addison-Wesley.

Mason, J. (1996). Expressing generality and roots of algebra. In: Bernarz, N., Kieran, C., Lee, L. (Eds). *Approaches to Algebra*. https://doi.org/10.1007/978-94-009-1732-3_5

Mason, J., Burton, L., & Stacey, K. (2010). *Thinking Mathematically* (2nd ed.). Pearson.

Mata-Pereira, J., & da Ponte, J. P. (2017). Enhancing students' mathematical reasoning in the classroom: teacher actions facilitating generalization and justification. *Educational Studies in Mathematics, 96*(2), 169-186. https://doi.org/10.1007/s10649-017-9773-4

Melhuish, K., Thanheiser, E., & Guyot, L. (2020). Elementary school teachers' noticing of essential mathematical reasoning forms: justification and generalization. *Journal of Mathematics Teacher Education, 23*, 35-67. https://doi.org/10.1007/s10857-018-9408-4

MoNE, (2013). *Ortaokul matematik dersi (5, 6, 7 ve 8. sınıflar) öğretim programı*.

MoNE, (2018). *Ortaokul matematik dersi (5, 6, 7 ve 8. sınıflar) öğretim programı*.

Mukuka, A., Balimuttajjo, S., & Mutarutinya, V. (2023). Teacher efforts towards the development of students' mathematical reasoning skills. *Heliyon, 9*(4). https://doi.org/10.1016/j.heliyon.2023.e14789

NCTM (2000). *Principles and Standards for School Mathematics*. NCTM.

Oflaz, G. (2017). *Determining the 8th graders ways of thinking and ways of understanding related to generalisation: A DNR based teaching experiment* (Publication No. 461560) [Doctoral dissertation, Gazi University]. Council of Higher Education Thesis Center.

Öz, T. (2017). *The examination of mathematical reasoning processes of the 7 th grade students* (Publication No. 475981) [Doctoral dissertation, Ataturk University]. Council of Higher Education Thesis Center.

Özmusul, B. (2018). *Investigation of mathematical justification skill levels of middle school 7th grade students* (Publication No. 548225) [Master's thesis, Gaziantep University]. Council of Higher Education Thesis Center.

Özmusul, B., & Bindak, R. (2022). Examining mathematical justification levels of 7th grade students. *Acta Didactica Napocensia*, *15*(2), 185-197. https://doi.org/10.24193/adn.15.2.12

Pedemonte, B. (2002). *Etude didactique et cognitive des rapports de l'argumentation et de la démonstration dans l'apprentissage des mathématiques* (Publication No. 00004579) [Doctoral dissertation, Genova University]. HALL Thesis.

Pedemonte, B. (2007). How can the relationship between argumentation and proof be analysed?. *Educational Studies in Mathematics, 66*(1), 23-41. https://doi.org/10.1007/s10649-006-9057-x

Pengmanee, S. (2016). Developing students' mathematical reasoning ability based on constructivist approach. *Journal of Advances in Humanities and Social Sciences*, *2*(4), 221-231. https://doi.org/10.20474/jahss-2.4.3

Peker, S. (2020). *Investigation secondary school students' generalization skills* (Publication No. 609365) [Master's thesis, Cumhuriyet University]. Council of Higher Education Thesis Center.

PISA (2022). PISA 2022 Turkey report. Ministry of National Education, General Directorate of Measurement, Evaluation and Examination Services. https://pisa.meb.gov.tr/meb_iys_dosyalar/2024_03/21120745_26152640_pisa2022_rapor.pdf

Reyes-Hernandez, L., & Mooney, E. (2021). *Examining middle school students' methods of justification*. https://ir.library.illinoisstate.edu/urs2021mat/3

Rodrigues, M., Brunheira, L., & Serrazina, L. (2021). A framework for prospective primary teachers' knowledge of mathematical reasoning processes. *International Journal of Educational Research*, *107*, Article 101750. https://doi.org/10.1016/j.ijer.2021.101750

Putra, A., Fauzi, K. M., & Landong, A. (2020). Differences in the improvement of students' mathematical reasoning ability and self-confidence between metacognitive approaches and realistic mathematical approaches in mts negeri balige. *Journal of Education and Practice*, *11* (36), 138-146. https://www.iiste.org/Journals/index.php/JEP/article/view/55142/56955

Säfström, A. I., Lithner, J., Palm, T., Palmberg, B., Sidenvall, J., Andersson, C., Boström, E., & Granberg, C. (2024). Developing a diagnostic framework for primary and secondary students' reasoning difficulties during mathematical problem solving. *Educational Studies in Mathematics*, *115*(2), 125-149. https://doi.org/10.1007/s10649-023-10278-1

Santos, L., Mata-Pereira, J., da Ponte, J. P., & Oliveira, H. (2022). Teachers' understanding of generalizing and justifying in a professional development course. *Eurasia Journal of Mathematics, Science and Technology Education*, *18*(1), 1-14. https://doi.org/10.29333/ejmste/11488

Serrazina, L., Brunheira, L., & Rodrigues, M. (2024). Developing elementary mathematics teachers' knowledge of mathematical reasoning processes. In *Thirteenth Congress of the European Society for Research in Mathematics Education (CERME13), Hungary,* https://hal.science/hal-04421415/document

Small, M. (2011). One, two, infinity. http://www.onetwoinfinity.ca/

Smit, R., Dober, H., Hess, K., Bachmann, P., & Birri, T. (2023). Supporting primary students' mathematical reasoning practice: the effects of formative feedback and the mediating role of self-efficacy. *Research in Mathematics Education, 25*(3), 277-300. https://doi.org/10.1080/14794802.2022.2062780

Staples, M., & Newton, J. (2016). Teachers' contextualization of argumentation in the mathematics classroom. *Theory into Practice, 55*(4), 294–301. https://doi.org/10.1080/00405841.2016.1208070

Stebbing, L. S. (1952). *A modern elementary logic.* University Publishers.

Stein, M. K., Engle, R. A., Smith, M., & Hughes, E. K. (2008). Orchestrating productive mathematical discussions: Five practices for helping teachers move beyond show and tell. *Mathematical Thinking and Learning*, *10*(4), 313–340. https://doi.org/10.1080/10986060802229675

Stylianides, A. J. (2007). The notion of proof in the context of elementary school mathematics. *Educational Studies in Mathematics, 65*(1), 1-20. https://doi.org/10.1007/s10649-006-9038-0

Stylianides, G. J. (2008). An analytic framework of reasoning-and-proving. *For the Learning of Mathematics, 28*(1), 9–16. https://www.jstor.org/stable/40248592

Staples, M. E., Bartlo, J., & Thanheiser, E. (2012). Justification as a teaching and learning practice: Its (potential) multifaceted role in middle grades mathematics classrooms. *The*

*Journal of Mathematical Behavior*, *31*(4), 447–462. https://doi.org/10.1016/j.jmathb.2012.07.001

Thanheiser, E., Melhuish, K., Sugimoto, A., Rosencrans, B., & Heaton, R. (2021). Networking frameworks: a method for analyzing the complexities of classroom cultures focusing on justifying. *Educational Studies in Mathematics*, *107*, 285-314. https://doi.org/10.1007/s10649-021-10026-3

Vale, C., Bragg, L. A., Widjaja, W., Herbert, S., & Loong, E. Y. K. (2017a). Children's mathematical reasoning: opportunities for developing understanding and creative thinking. *Australian Primary Mathematics Classroom*, *22*(1), 3-8. https://search.informit.org/doi/abs/10.3316/informit.735199590694733

Vale, C., Widjaja, W., Herbert, S., Bragg, L. A., & Loong, E.Y. K. (2017b). Mapping Variation in Children's Mathematical Reasoning: The Case of 'What Else Belongs?'. *International Journal of Science and Mathematics Education*, *15*, 873–894. https://doi.org/10.1007/s10763-016-9725-y

Venenciano, L., & Heck, R. (2016). Proposing and testing a model to explain traits of algebra preparedness. *Educational Studies in Mathematics*, *92*, 21-35. https://doi.org/10.1007/s10649-015-9672-5

Visnovska, J., & Cobb, P. (2009). Learning about building mathematics instruction from students' reasoning: a professional development study. In Hunter, R., Bicknell,B., & Burgess, T. (Eds.), *Proceedings of the 32nd annual conference of the Mathematics Education Research Group of Australasia* (pp. 547-554).

https://www.researchgate.net/profile/Jana-Visnovska/publication/43515267_Learning_about_building_mathematics_instruction_from_students'_reasoning_A_professional_development_study/links/0deec528afe9ebf661000000/Learning-about-building-mathematics-instruction-from-students-reasoning-A-professional-development-study.pdf

Waluyo, S., Armanto, D., & Mansyur, A. (2021). Analysis of mathematic reasoning ability and self-confidence students using models pbl reviewed from intelligence intrapersonal in smpn 5 youth vocationality. *Journal of Education and Practice*, *12*(16), 59-64. https://doi.org/10.7176/JEP/12-16-08

Watson, F. R. (1980). The role of proof and conjecture in mathematics and mathematics teaching. *International Journal of Mathematical Educational in Science and Technology*, *11*(2), 163-167. https://doi.org/10.1080/0020739800110202

Widjaja, W., Vale, C., Herbert, S., Loong, E. Y., & Bragg, L. A. (2021). Linking comparing and contrasting, generalizing and justifying: a case study of primary students' levels of justifying. *Mathematics Education Research Journal*, *33*, 321-343. https://doi.org/10.1007/s13394-019-00306-w

Widjaja, W., & Vale, C. (2021). Counterexamples: Challenges faced by elementary students when testing a conjecture about the relationship between perimeter and area. *Journal on Mathematics Education, 12*(3), 487-506.https://doi.org/10.22342/jme.12.3.14526.487-506

Yeşildere, S., & Türnüklü, E. B. (2007). Examination of students' mathematical thinking and reasoning processes. *Ankara University Journal of Faculty of Educational Sciences*, *40*(1), 181-213. https://doi.org/10.1501/Egifak_0000000156

Yin, R. K. (2009). *Case study research: Design and methods* (4th ed.). Sage.

Necatibey Faculty of Education Electronic Journal of Science and Mathematics Education
Vol. 18, Issue 2, December 2024, pp. 324-350 ISSN: 1307-6086

Research Article

# Quantitative and Qualitative Comparison of Learning Outcomes in Mathematics Curricula of Different Countries in the Context of Mathematical Literacy Skills

**İrem ÇEVİK[1], Bilgen KERKEZ[2]**

[1] Ministry of National Education, Musalı Veli Haşim Çiftçi Middle School, Türkiye, iremuz33@gmail.com, http://orcid.org/0009-0004-5710-6735

[2] Ministry of National Education, General Directorate of Secondary Education, Türkiye, bilgenkerkez@gmail.com, http://orcid.org/0000-0002-5124-6030

Received : 04.06.2024 Accepted : 12.09.2024

Doi: https://doi.org/10.17522/balikesirnef.1495776

---

*Abstract* – In this qualitative study, written resources related to the curricula of the countries and 2022 PISA results were used to examine the learning outcomes in the mathematics curricula of different countries in quantitative and qualitative dimensions. In the quantitative dimension, other countries had higher PISA scores and rankings with fewer learning outcomes than Türkiye. In the qualitative dimension, it was observed that the outcomes were concentrated in the mathematical process of third mathematical literacy in Türkiye and Poland, and in the mathematical process of second mathematical literacy in Sweden and Denmark. Outcomes in Türkiye were found to be more complex and not evenly distributed across mathematical processes. Based on the findings, it is suggested that the mathematics learning outcomes should be reduced and distributed more evenly across the mathematical processes, and a renewed programme should be proposed to enable in-depth development of skills as well as knowledge.

*Key words*: Mathematical literacy skills, mathematical processes, mathematics curricula, learning outcomes.

------------

Corresponding author: İrem ÇEVİK, iremuz33@gmail.com

## Introduction

In order to keep pace with the continuous change and development in mathematics education worldwide, it is inevitable to focus on improvements in mathematics teaching. As part of the ongoing effort to improve mathematics education, mathematics curricula need to be regularly revised. Although research on the comparison of curricula for the renewal of curricula is increasing both in Türkiye and worldwide (Bal İncebacak, 2022; Chung & Chung, 2008; Deveci, 2018; Güven & Gürdal, 2011; Öçal, 2017; Revina & Leung, 2018; Yang et al., 2017; Yavuz Topaloğlu & Balkan Kıyıcı, 2015; Xie & Carspecken, 2019), studies comparing the curricula of Türkiye and other countries generally focus on the similarities and differences of some or all of the curricula elements consisting of objectives, content, learning-teaching process and assessment and evaluation approaches (Erbilge, 2019; Çetinbağ, 2019; Çoban, 2011; Güzel et.al., 2010; Karataşlı, 2019; Yazıcıoğlu & Pektaş, 2018). In the process of curricula renewal, the learning outcomes to be gained with the target element, which is one of the basic components of the curricula, also have an important place. The fact that the changes made in the number of learning outcomes in the curricula in many curriculum renewals carried out many times in our country attract attention (Çobanoğlu & Yıldırım, 2021) suggests that the quantity of learning outcomes that meet the content of the curricula is important and is the first step in changing the quality of the curricula. In addition, there are study results in the literature that reveal the importance of the number of learning outcomes such as the fact that teaching leaves education behind due to the intensity of the programmes (Kahramanoğlu et al., 2016), that there are intensive programme contents (Çaycı, 2018), and that the programmes have not reached the desired level in terms of gaining the skills to be used in real life (Kaya & Karakaya, 2012). In this context, this research is directed as a quantitative comparison of the learning outcomes in the mathematics curricula of our country and different countries.

When it comes to the renewal of mathematics curricula, one of the important criteria taken into consideration is the international examinations that enable the determination of the leading countries in achievement. The comparison of different countries with international exams provides insights for countries in the revision, development and renewal of mathematics curricula. In this respect, the large-scale international exams administered by the Programme for International Student Assessment (PISA) provide countries with the opportunity to see their global rankings in terms of achievement in mathematics performance and their mathematical literacy levels, and to understand the literacy dimension in the curricula and the literacy skills reflected by the curricula elements. Since increasing mathematical literacy is accepted as a common goal when countries design their programmes, this study aims to examine the mathematical processes of mathematical literacy skills in the programmes of the countries that are successful in the 2022 PISA exam in achieving this goal and to identify both similarities and differences by comparing them with the Turkish programme. This comparison will contribute to the literature by providing valuable information in the following sections of the study.

Among the 81 countries that participated in the 2022 PISA, Denmark and Poland are among the countries that stand out, ranking 13$^{th}$ and 15$^{th}$ in mathematics performance, respectively. Sweden, another selected country, ranks 22nd and demonstrates remarkable success by ranking among the top 25 countries in terms of 2022 PISA mathematics performance. However, Türkiye ranks 39$^{th}$ in terms of mathematics performance. In addition, 2022 PISA defines eight proficiency levels from the lowest to the highest: 1a, 1b, 1c, 2, 3, 4, 5 and 6 (MoNE, 2023). In terms of proficiency level, Türkiye is classified at Level 2, whereas the other three countries in the study are at Level 3. It is thought that the mathematical literacy skills reflected in the programmes play a role behind the proficiency levels and successful rankings in PISA. In this sense, it will be important to make comparisons by considering the target element of the programmes of these countries, which are selected because they are more successful than Türkiye in terms of both proficiency level and ranking, and to see to what extent mathematical literacy skills are included in mathematics lesson outcomes in order to determine the missing or defective situations in the curricula. When the literature is examined, it is seen that studies on the promotion and development of mathematical literacy focus on improving students' mathematical literacy skills through the creation of appropriate learning environments (Çilingir Altıner & Dinç Artut, 2017; Çilingir & Dinç Artut, 2016; Dibek et al., 2016; Gürbüz, 2014; Karakaş & Ezentaş, 2021; Köysüren & Üzel, 2018; Var & Altun, 2021). It has been realized that there is a need for research that focuses on the extent to which and in which mathematical processes mathematical literacy is included by addressing the goal element of the curricula. In this context, this study aims to compare the objectives in the mathematics curricula of Poland, Sweden, Denmark and Türkiye in terms of the mathematical processes of mathematical literacy, to examine their similarities and differences, and to determine the qualitative situation in the context of mathematical literacy skills. Thus, this study, which focuses on the subject, will provide an important reference point for future research.

The 21st century modern society needs not only content knowledge but also skills such as critical thinking, problem solving, decision making, creativity, innovation, questioning and reasoning skills. Mathematical literacy is one of the necessary components for the development of 21st century skills. In this context, the student, who should have the opportunity to directly experience these skills by building his/her own literacy, needs an environment where curricula are implemented in which mathematical literacy skills are structured as a whole by focusing on all processes. In the 2022 PISA assessment, mathematical literacy is the capacity of an individual to reason mathematically, formulate, use and interpret mathematics in order to solve various problems in real life. 2022 PISA measures how effectively countries prepare their students to use mathematics in all areas of their lives, including professional fields, as individuals equipped with 21st century skills (MoNE, 2023). The processes of 2022 PISA-defined mathematical literacy skills are outlined in Table 1.

**Table 1** Processes of Mathematical Literacy Skills

| Mathematical reasoning | | |
|---|---|---|
| Formularisation | Utilisation | Interpretation and evaluation |
| -To be able to realise mathematical concepts<br><br>-To be able to gain mathematical structure for a problem<br><br>-Formulate in mathematical terms | -To be able to do arithmetic calculations, to solve equations, to make logical inferences based on assumptions, to make symbolic arrangements<br><br>-To be able to extract mathematical information using tables and graphs, to be able to show and organise spatial shapes, to be able to analyse data | -To be able to transform mathematical solutions or reasoning into problem context<br><br>-To be able to determine whether the mathematical results are reasonable and whether they make sense in the context of the problem |

The mathematical reasoning process, which is the first of the mathematical literacy processes as outlined in Table 1, is the main process that covers the other processes. This process is the capacity to use mathematical concepts, tools and logic to conceptualise real-life problems or situations and to generate solutions for these problems or situations. Students who are mathematically literate in the process of mathematically formulating problems or situations have the ability to recognise mathematical concepts and ideas in the problems they encounter and to provide a mathematical structure to these problems. Students who are mathematically literate in using mathematical concepts, facts and processes in the third process have the ability to solve mathematically formulated problems by using appropriate mathematical tools to achieve mathematical results. In the last process, students who are mathematically literate in interpreting, applying and evaluating mathematical outputs have the ability to reflect on mathematical solutions, results or inferences and interpret them in the context of real life problems that initiated the process (MoNE, 2023). In order to raise mathematically literate individuals who can cope with daily problems and to end the unfamiliarity with mathematical literacy as soon as possible, it is necessary to integrate the four processes mentioned in the curricula. Based on this necessity, this study is important in terms of showing the extent to which the learning outcomes in the mathematics curricula of the identified countries demonstrate mathematical literacy including the specified mathematical processes and will provide important inputs to the literature in terms of focusing on showing which processes of mathematical literacy the learning outcomes are equipped with. In this context, the study aims to investigate the reasons for the differences in mathematical literacy skills between Türkiye and Poland, Sweden and Denmark in terms of learning outcomes in mathematics curricula.

As a result, this research is a guide for addressing the learning outcome dimension by prioritising meaningful and in-depth learning in curricula and taking concrete steps to simplify and deepen the content. On the other hand, using the informative and guiding nature of PISA, which helps to predict the weight given to mathematical literacy by mathematics curricula on a global scale, the fact that the learning outcomes of successful countries will be compared with the learning outcomes of our programme for the purpose of raising mathematically literate individuals has the quality of shedding light on the curricula development process in order to develop and effectively use mathematical literacy skills. Based on the idea that the quantity and quality of learning outcomes have a role in raising individuals who have acquired literacy skills, it is thought that studies should be carried out to use the appropriate quantity, meaningful and deepening knowledge structures instead of too much information in the outcomes and to integrate essential skills, such as literacy, with the knowledge structure. In this study, it is aimed to quantitatively examine the learning outcomes in the mathematics curricula implemented in the second level of compulsory education in Poland, Sweden, Denmark and Türkiye and their performances according to 2022 PISA mathematics performances and mathematical literacy skills processes and to classify the learning outcomes in these programmes according to mathematical literacy skills processes. Within the scope of this main purpose, the following questions were answered.

1) How do Poland, Sweden, Denmark and Türkiye's mathematics learning outcomes, 2022 PISA mathematics performance and mathematical processes of mathematical literacy skills vary?

2) How are the mathematics learning outcomes of Poland, Sweden, Denmark and Türkiye classified according to the mathematical processes of the 2022 PISA mathematical literacy skills?

## Method

The study, which aims to compare the learning outcomes of mathematics lessons in Poland, Sweden, Denmark and Türkiye quantitatively and qualitatively in terms of mathematical literacy skills, was carried out in order to reveal an existing situation. In this section of the study, research design, data collection, data analysis and validity and reliability are presented. In order to ensure conceptual integrity, the term "learning outcome" was used in this study instead of different expressions such as objective, outcome, etc. used in the selected countries to refer to the first element of the programme.

### Research Design

This research was carried out using the document review approach, which is one of the qualitative research methods, to achieve the determined purpose. Document review is an approach that is effective in filling the gaps in the literature and enriching the knowledge in the research field based on the information obtained by in-depth analysis of written sources related to the research field (Creswell, 2017). In this context, the learning outcomes of the current mathematics curricula and the 2022 PISA mathematics literacy scores of Poland, Sweden, Denmark and Türkiye were examined in this study. The study analysed the similarities and differences in the mathematical literacy processes reflected in the learning outcomes of the mathematics curricula of these countries.

### Data Collection

The data collection tools include the current mathematics curricula documents obtained from the websites of official institutions such as ministries of education, universities, and national education institutions of the identified countries, the 2018 mathematics curricula implemented in Türkiye, and the 2022 PISA Türkiye report published by the Ministry of National Education.

## Data Analysis

The first step to start the data analysis process involved analysing the learning outcomes for mathematics in the current curricula of the selected countries and Türkiye. In this step, numerical data on mathematics learning outcomes for each country and Türkiye were collected and changes in the number of learning outcomes were analysed. Then, each country's 2022 PISA mathematics performances, their performances in mathematical literacy processes and their mathematical literacy levels were carefully analysed and compared to assess the general situation of the countries. Finally, a comprehensive evaluation was made to determine how the learning outcomes in the mathematics curricula of the countries were classified according to the 2022 PISA mathematical literacy processes and to identify similarities and differences in the classification. In this section, descriptive analysis technique is the preferred method for data analysis. This method allows the data to be systematically defined and organized based on certain themes (Yıldırım & Şimşek, 2008). In the context of 2022 PISA, formularisation, utilisation and interpretation mathematical processes constitute the themes of this study. The learning outcomes presented by the curricula of the identified countries were analyzed systematically in line with these themes. These analyses revealed how the curricula of different countries develop mathematical literacy skills and in which processes this skill is strong or weak. Finally, the findings of this analysis are meticulously described, presented in detailed tables and interpreted.

## Validity and Reliability

In order to evaluate the reliability of the collected documents, the authors of the documents, the publication dates, the reliability of the publishing websites and the purpose of the documents were examined comprehensively. In order to ensure validity, the research process was planned in detail by the researchers, the agreement between the researchers was compared for the data obtained, and the findings obtained after data analysis were supported with numerical data and examples.

# Findings and Discussions

The findings, which are presented in tables after analysing the data collected for the purpose of the research, consist of two sections to answer the two sub-questions of the research. In the first part, quantitative findings related to mathematics lesson learning outcomes and mathematical literacy skills are discussed. In the second part, the qualitative dimension of the learning outcomes in mathematics curricula according to mathematical literacy processes is discussed.

## Quantitative Findings on Mathematics Lesson Learning Outcome and Mathematical Literacy Skills

### *Number of Mathematics Learning Outcomes of Poland*

In 2017, Poland undertook a comprehensive reform of its education system, to align it with the demands of the 21st century. As part of this reform, it extended the duration of compulsory education to 9 years, including pre-school education, by increasing the duration of compulsory education to 8 years. After compulsory education, students are offered a four-year general upper secondary education (Podstawa, 2018; Wojnak & Majorek, 2018). The primary and secondary school programme is divided into two phases under the reform. The first phase covers grades 1-3and aims to enable students to acquire basic skills and knowledge. The second phase, covering grades 4-8, aims for students to develop their knowledge and skills in more depth (Podstawa, 2017). The core programme for each lesson includes specific objectives and learning outcomes that students should achieve at the end of the lessons. The learning outcomes and learning areas determined for the mathematics lesson are presented in Table 2 for grades 4-6th.

**Table 2** Poland 4-6th Grades Mathematics Lesson Learning Outcomes and Learning Areas

| Learning areas | Number of learning outcomes |
|---|---|
| Numbers | 54 |
| Algebra | 2 |
| Geometry | 36 |
| Statistics | 2 |
| Problem solving | 7 |
| Total | 101 |

Table 2 indicates that the Polish mathematics curricula for grades 4-6 includes 5 learning domains and 101 learning outcomes. Among these learning domains, the highest number of learning outcomes is in the learning domain of numbers. The lowest number of learning outcomes is in algebra and statistics learning areas.

### *Number of Mathematics Learning Outcomes of Sweden*

Sweden changed its curricula in 2022, updating the curricula that had been in use since 2011 for all grade levels within compulsory education. The new curricula aims to simplify the learning outcomes and to increase students' participation in the learning process and teachers' flexibility (Lidbäck, 2021). Compulsory education applies to all children between the ages of 6-16 and consists of four stages: pre-school, primary school, middle school and high school (Skolverket, 2022). The learning outcomes and learning areas for grades 4-6 of the Swedish mathematics curricula are presented in Table 3.

**Table 3** *S*weden 4-6th Grades Mathematics Lesson Learning Outcomes and Learning Areas

| Learning areas | Number of learning outcomes |
|---|---|
| Numbers | 8 |
| Algebra | 5 |
| Geometry | 5 |
| Probability and statistics | 4 |
| Relationships and change | 3 |
| Problem solving | 2 |
| Total | 27 |

Table 3 shows that there are 6 learning domains and 27 learning outcomes in the Swedish mathematics curricula for grades 4-6. Among these learning domains, the highest number of learning outcomes is in the domain of numbers, while the lowest number of learning outcomes is in the domain of problem solving.

***Number of Mathematics Learning Outcomes of Denmark***

In Denmark, education is offered as compulsory education in grades 1-9 and optional education in grade 10 and is based on the national framework programme. Mathematics curricula start from grade 1 and continue until grade 10. At primary school level, the mathematics curricula for 2019 are organised into grades 1-3, grades 4-6 and grades 7-9 (Danmarks Læringsportal [EMU], 2019). The number of learning outcomes and learning areas of mathematics curricula in grades 4-6 are presented in Table 4.

**Table 4** Denmark 4-6th Grades Mathematics Lesson Learning Outcomes and Learning Areas

| Learning areas | Number of learning outcomes |
|---|---|
| Mathematical competences | 26 |
| Numbers and algebra | 18 |
| Geometry and measurement | 24 |
| Statistics and probability | 12 |
| Total | 80 |

According to Table 4, there are 4 learning domains and 80 learning outcomes in the Danish mathematics curricula for grades 4-6. Among these learning areas, the area with the highest number of learning outcomes is mathematical competences and the area with the lowest number of learning outcomes is statistics and probability.

***Number of Mathematics Learning Outcomes of Türkiye***

Türkiye has a 12-year compulsory education system. Mathematics curricula start from grade 1 and continue through grade 12. Mathematics curricula of 2018 are used at primary, middle and high school levels. The learning outcomes and learning areas of mathematics curricula for grades 5-8 are presented in Table 5.

**Table 5** Türkiye 5-8th Grades Mathematics Lesson Learning Outcomes and Learning Areas

| Learning areas | Number of learning outcomes |
|---|---|
| Numbers and operations | 106 |
| Algebra | 23 |
| Geometry and measurement | 67 |
| Data processing | 14 |
| Probability | 5 |
| Total | 215 |

According to Table 5, there are 5 learning areas and 215 objectives in the current mathematics curricula. In the current mathematics curricula, the order from the highest number of objectives to the lowest number of objectives is numbers and operations, geometry and measurement, algebra, data processing and probability.

As a result, the total number of learning outcomes of the second level mathematics curricula of the countries are listed as Türkiye, Poland, Sweden and Denmark, respectively. The finding that Türkiye's mathematics curricula has the highest number of learning outcomes compared to the curricula of other countries in the total number of learning outcomes of the countries suggests that the content of the Turkish mathematics curricula is more intense compared to other countries.

**Performance of Countries in 2022 PISA Mathematical Literacy Skills**

The mathematical literacy performances and proficiency levels of Poland, Sweden, Denmark and Türkiye according to the 2022 PISA results and their performances according to mathematical processes are presented in Table 6.

**Table 6** Countries' 2022 PISA Mathematical Literacy and Mathematical Processes Scores and Proficiency Levels

| Country | Mathematics literacy score | Mathematics literacy proficiency level | Mathematical processes | | |
|---|---|---|---|---|---|
| | | | Formularisation | Utilisation | Interpretation |
| Denmark | 489 | 3 | 485 | 488 | 491 |
| Poland | 489 | 3 | 485 | 491 | 490 |
| Sweden | 482 | 3 | 474 | 481 | 478 |
| Türkiye | 453 | 2 | 451 | 452 | 455 |

As seen in Table 6, when the scores of the countries are analysed, it is seen that Türkiye's mathematical literacy performance is lower than the performance of other countries. In addition, while Türkiye's mathematical literacy proficiency level was 2, the other countries were found to be at the 3rd proficiency level. When the scores according to mathematical processes are analysed, it is seen that the performances in each process in Türkiye are considerably lower than the performances of other countries.

## Qualitative Findings of Mathematics Lesson Learning Outcomes in the Context of Mathematical Literacy Skills

Mathematical reasoning, which is one of the processes needed to achieve mathematical literacy, was not used in the classification since it covers the other three processes. The processes of mathematical literacy skills included in the classification are formulating situations mathematically, using mathematical concepts, facts and processes, and interpreting, applying and evaluating mathematical outputs. The classification of the learning outcomes in the learning domains of the mathematics curricula of the countries according to these processes is given in tables for each country. In the tables, the number of learning outcomes in each learning domain, which reflects which process predominantly, was written in the relevant section of the table of the country and the classification of the existing learning outcomes was made.

### *Poland*

The classification of the learning outcomes in the learning areas of the Polish mathematics curricula according to the mathematical literacy skill processes is given in Table 7.

**Table 7** Distribution of Poland's Mathematics Learning Outcomes in Mathematical Processes

| Learning areas | Mathematical processes | | |
|---|---|---|---|
| | Formularisation process | Utilisation process | Interpretation process |
| Numbers | 9 | 40 | 5 |
| Algebra | 2 | - | - |
| Geometry and measurement | 11 | 17 | 8 |
| Statistics | - | 2 | - |
| Problem solving | 2 | 1 | 4 |

As seen in Table 7, 24 learning outcomes of the Polish mathematics lesson are concentrated in the 2nd process, 60 learning outcomes are concentrated in the 3rd process and 17 learning outcomes are concentrated in the 4th process. In this case, it can be concluded that the learning outcomes are most concentrated in the 3rd process and least in the 4th process. As an example of learning outcomes aligned with mathematical processes, the outcome "Uses letter notation for unknown numerical quantities and writes simple algebraic expressions based on embedded knowledge in a practical context" is aligned with the formularisation process. Another example learning outcome, "Draws angles less than 180°." is associated with the process of utilisation. In addition, the learning outcome "Verifies the result of a text task." is an example of an outcome appropriate for the interpretation process. Although the number of outputs in the 4th mathematical process is the lowest, the fact that it is close to the number of outputs in the 2nd process shows that its amount is not to be underestimated. When we look at the distribution of mathematical processes according to learning areas, it is seen that all mathematical processes are mentioned in the areas of numbers, geometry and measurement and problem solving. In the learning domains of algebra and statistics, it is seen that the number of outcomes is low and there are mathematical processes that are not used in learning outcomes**.** It was determined that the mathematical processes of utilisation and interpretation were not used in the algebra learning area and the mathematical processes of formularisation and interpretation were not used in the statistics learning area.

When the structure of the learning outcomes in the Polish mathematics curricula is examined, it is seen that in most of the outcomes, the processes are included separately rather than intertwined. For example, it can be said that the outcome "Separates two-digit numbers into their prime factors" belonging to the learning domain of numbers is only for the 3rd process.

### *Sweden*

The classification of the learning outcomes in the learning areas of the Swedish mathematics curricula according to mathematical literacy processes is given in Table 8.

**Table 8** Distribution of Sweden's Mathematics Learning Outcomes in Mathematical Processes

| Learning areas | Mathematical processes | | |
|---|---|---|---|
| | Formularisation process | Utilisation process | Interpretation process |
| Numbers | 5 | 2 | 1 |
| Algebra | 3 | 2 | - |
| Geometry | 2 | 1 | 2 |
| Probability and statistics | 1 | 2 | 1 |
| Relationships and change | 2 | 1 | - |
| Problem solving | 1 | - | 1 |

According to Table 8, it is seen that 14 of the 27 learning outcomes of the Swedish mathematics lesson are concentrated in the 2nd process, 8 in the 3rd process and 5 in the 4th process. In this case, it is seen that the learning outcomes are most concentrated in the 2nd process and least in the 4th process. As an example of learning outcomes aligned with mathematical processes, the outcome "Variables and their construction in simple algebraic expressions and equations" is aligned with the formularisation process. Another example of a learning outcome is "Using digital tools in calculations", which is associated with the utilisation process. In addition, the learning outcome "Evaluating the plausibility of predictions and calculations" is an example of an outcome appropriate for the interpretation process. However, although the distribution of outputs in the 3rd and 4th processes is low, since the total number of outputs is also low, it is seen that they are not negligible compared to the total number of outputs. When the distribution of mathematical processes according to learning areas is analysed, it is seen that all processes are mentioned in the learning areas of numbers, geometry, probability and statistics. In the learning outcomes in algebra, relationships and change and problem solving learning areas, it is determined that there are mathematical processes that are not used**.** It was determined that the mathematical process of interpretation was not used in algebra, relations and change learning areas and the mathematical process of utilisation was not used in problem solving learning area.

When the structure of the learning outcomes in the Swedish mathematics curricula is analysed, it is seen that in a few of the outcomes, process components and topics are intertwined rather than separately. For example, the outcome "Has knowledge about rational numbers, including negative numbers, and their properties, and how they can be divided and used" belonging to the learning domain of numbers is utilised for both process 2nd and process 3rd.

***Denmark***

The classification of the learning outcomes in the learning areas of the Danish mathematics curricula according to mathematical literacy processes is given in Table 9.

**Table 9** Distribution of Denmark's Mathematics Learning Outcomes in Mathematical Processes

| Learning areas | Mathematical processes | | |
|---|---|---|---|
| | Formularisation process | Utilisation process | Interpretation process |
| Mathematical competences | 13 | 10 | 3 |
| Numbers and algebra | 9 | 7 | 2 |
| Geometry and measurement | 13 | 8 | 3 |
| Statistics and probability | 6 | 4 | 2 |

When Table 9 is analysed, it is seen that 41 of the learning outcomes of the mathematics lesson in Denmark are concentrated in the 2nd process, 29 in the 3rd process and 10 in the 4th process.  As an example of learning outcomes aligned with mathematical processes, the outcome "The student can formulate simple algebraic expressions for calculations." is aligned with the formularisation process. Another example learning outcome, "The student can translate between everyday language and expressions using mathematical symbols" is associated with the process of utilisation. In addition, the learning outcome "The student can estimate and determine the perimeter and area." is an example of an outcome suitable for the interpretation process. In this case, it is seen that the learning outcomes are concentrated mostly in the 2nd process and least in the 4th process. It was determined that all processes were mentioned in the learning outcomes of the mathematics lesson. When the distribution of mathematical processes according to learning areas is analysed, it is seen that mathematical competencies, numbers and algebra, geometry and measurement, statistics and probability are mentioned in all areas.

When the structure of the learning outcomes in the Danish mathematics curricula is analysed, it is seen that most of the outcomes are not intertwined with the process components, but they are separated. For example, it can be said that the outcome "The student can use simple mathematical models" belonging to the learning domain of mathematical competences is only related to the 3rd process.

### *Türkiye*

The classification of the learning outcomes in the learning areas of the 2018 mathematics curricula used at the second level in Türkiye according to mathematical literacy processes is given in Table 10.

**Table 10** Distribution of Türkiye's Mathematics Learning Outcomes in Mathematical Processes

| Learning areas | Mathematical processes | | |
|---|---|---|---|
| | Formularisation process | Utilisation process | Interpretation process |
| Numbers and operations | 30 | 70 | 6 |
| Algebra | 10 | 13 | - |
| Geometry and measurement | 27 | 39 | 1 |
| Data processing | 2 | 12 | - |
| Probability | 2 | 3 | - |

When Table 10 is analysed, it is seen that 137 learning outcomes are concentrated in the 3rd mathematical process in the mathematics lesson in Türkiye. 71 learning outcomes are concentrated in the 2nd mathematical process and the concentration in the 4th mathematical process is the lowest with 7 outcomes. It was determined that the learning outcomes of the mathematics lesson were not equally distributed among mathematical processes. When the distribution of mathematical processes according to learning areas is analysed, it is seen that all of the processes are mentioned in the areas of numbers and operations, geometry and measurement; but there are no learning outcomes for the 4th process in the learning areas of algebra, data processing and probability.  As an example of learning outcomes that are compatible with mathematical processes, the outcome "Recognizes the equation with a first-order unknown and constructs an equation with a first-order unknown in accordance with the given real-life situations." is compatible with the formularisation process. Another sample learning outcome, "Expresses rational numbers in decimal notation." is associated with the process of utilisation. In addition, the learning outcome "Predicts the result of operations with fractions." is an example output suitable for the interpretation process.

When the structure of learning outcomes is analysed, it is seen that most of the learning outcomes in Türkiye are structured in a way to cover more than one process. For example, the outcome "Forms the surface area relation of a right circular cylinder; solves related problems." belonging to the 8th grade geometry learning domain is structured for both the 2nd and 3rd process.

# Conclusions and Suggestions

## Conclusions Related to The Quantitative Dimension

According to the findings of the quantitatively analysed learning outcomes, the number of learning outcomes for Sweden's mathematics curricula is 27, Denmark's learning outcomes is 80 and Poland's learning outcomes is 101, while the number of learning outcomes for Türkiye's mathematics curricula is 215. It is noteworthy that the number of learning outcomes of the Danish and Swedish mathematics curricula is quite low compared to Türkiye. In support of this result, Duygu (2013) and Çoban (2011) state that Türkiye's mathematics curricula has a higher number of learning outcomes compared to the other countries compared. When compared according to learning domains, the learning outcomes in learning domains such as "Numbers and Operations", "Algebra", "Geometry and Measurement", "Data Processing" and "Probability" in Türkiye's curricula are higher than the numerical values of learning outcomes in similar learning domains of other countries.

2022 PISA Mathematics literacy performances show that Sweden's score is 482, Denmark and Poland's score is 489. Türkiye's mathematical literacy score is 452. With these scores, it is seen that Poland, Sweden and Denmark are at the 3rd mathematical literacy proficiency level and Türkiye is at the 2nd mathematical literacy proficiency level. This suggests that there is a relationship between the number of learning outcomes and mathematical literacy success. This idea is consistent with the fact that having too many learning outcomes in a programme may have a negative effect on the depth and effectiveness of teaching (Hook et al., 2007; Schoen et al., 2011). As a matter of fact, according to Özgün-Koca and Şen (2002), the content density of the programmes has a negative effect on the development of mathematical literacy, and according to Buluş Kırıkkaya (2009), there are teacher opinions that there is a need to simplify the content and the number of learning outcomes in order to increase achievement. Similarly, according to Hobson (2001), one of the biggest obstacles for students to become mathematically literate is the high level of dense content in curricula. In addition, it is also stated in the literature that in programmes with a high number of learning outcomes, teachers have difficulty in completing the learning outcomes in the prescribed lesson times (Akpınar, 2004; Çaycı, 2018), they cannot use methods and techniques that will increase the active participation of students due to the concern of the programme (Ayvacı & Durmuş, 2013), and therefore the achievement of students who experience a decrease in their motivation (Dursun & Dede, 2004). In the 2022 PISA assessment, the fact that the countries that were successful but not included in the study also had a low number of learning outcomes supports the findings of the study. For example, it is seen that Singapore ranked first in the 2022 PISA with the highest scores in mathematics performance and mathematical literacy processes, as in almost all of the international exams it participated in. In the study by Erdoğan et al. (2016) comparing the mathematics curricula of Türkiye and Singapore, it was stated that there were fewer learning outcomes in the mathematics curricula implemented in Singapore. Based on this result, it can be concluded that reducing the number of learning outcomes in the curricula will positively affect mathematics literacy.

In general terms, based on the finding that the mathematics curricula of the countries ranked higher in PISA are generally simpler and the number of learning outcomes is lower, it is recommended to make changes in the curricula of the mathematics lesson being implemented in Türkiye in order to deepen the content and to reduce the number of learning outcomes to a great extent. The significant decrease in the number of learning outcomes in the renewed curricula will pave the way for the development of mathematical literacy skills by further strengthening the educational approach that aims to build knowledge together with skills and the belief that students should focus on holistic development.

### Conclusions Related to The Qualitative Dimension

When the learning outcomes are analysed qualitatively, it is seen that the reflections in the outcomes of the programmes designed to develop mathematical literacy differ. When the distribution of the learning outcomes in the programmes according to the mathematical processes of mathematical literacy is examined, it is seen that Poland's mathematics curricula has more learning outcomes in the process of using mathematical concepts, facts and processes, while Sweden and Denmark's curricula has more learning outcomes in the process of formulating situations mathematically. However, in these countries, other mathematical processes are also addressed to a sufficient extent and the processes are distributed to the outcomes in a balanced way. In Türkiye's mathematics curricula, there are more learning outcomes in the process of using mathematical concepts, facts and processes. This result suggests that individuals educated in Türkiye may have more mathematical literacy capacity in using mathematical concepts, facts and processes. In addition, it is seen that there are sufficient learning outcomes in the process of formulating situations mathematically, which is the second process; however, there are not enough learning outcomes that meet the fourth process. Based on this result, it can be said that not all mathematical literacy processes are sufficiently emphasised in mathematics lesson learning outcomes and the distribution is not balanced. In the programmes of the three successful countries, learning outcomes of mathematical literacy skills were balanced in all processes, whereas in Türkiye, it was found that learning outcomes had an unbalanced distribution in terms of mathematical literacy skill processes. This situation suggests that there is a relationship between ensuring a balanced distribution of mathematical processes to outcomes and mathematical literacy success. As a matter of fact, a student with developed mathematical literacy benefits from mathematics at the highest level (Altun, 2015). It is thought that the PISA results applied on a global scale reflect this situation (MoNE, 2019; MoNE, 2023). This situation suggests that these three successful countries, which stand out with PISA, give priority to the development of mathematical literacy skills holistically while designing their curricula, and focus on the depth of skill development with its processes by going beyond traditional knowledge transfer. Based on this idea, it is considered important that mathematical literacy processes, which start with mathematical reasoning and continue with formulating, using and interpreting mathematics, should be handled in a holistic manner for a more effective mathematics teaching and mathematics success. In this respect, an integrated structure can be created in the curricula by placing the mathematical literacy skills and processes, which are necessary to include and use the skills needed in the 21st century in daily life, more balanced in the learning outcomes of the curricula and in the lesson contents that meet these outcomes. Aware that the practice and development of 21st century skills are intertwined with literacy skills, the mathematics teaching process is shaped around the idea that these skills are interrelated and necessary for use in daily life. A new curricula proposal that covers all processes of mathematical literacy in a balanced way will be a valuable resource for revising not only mathematics but also curricula in various other fields.

When the distribution of the learning outcomes in each mathematics learning domain in terms of mathematical processes is analysed, it is seen that in all of the countries participating in the study, including Türkiye, there are missing mathematical processes in some learning domains. This is thought to be due to the fact that the learning outcomes within the scope of learning areas tend to provide knowledge such as numbers, algebra, geometry and statistics rather than mathematical literacy. There is a need to include learning outcomes that adequately address all mathematical processes both in general and within the scope of each learning area in the curricula. Thus, it would be an important initiative to organise the outcomes or include new outcomes in each learning area of mathematics in a way to enable students to experience each process of mathematical literacy skills.

Although the mathematical process emphasised in the Turkish curricula is the same as in Poland, the reason for the difference in the achievement of Poland and Türkiye may be that most of the learning outcomes in Türkiye are structured in such a way that they cover more than one literacy process. The fact that the outcomes have a complex structure resulting from the intertwining of the mathematical processes that constitute the skill may make it necessary to follow teaching and learning practices in terms of mathematical literacy. As a matter of fact, it is also important to ensure the integration of literacy into the classroom (Steinberg, 2011) in the development of literacy skills. In this case, in order to gain mathematical literacy skills in the classroom, all the processes covered by the learning outcome must be fulfilled. In teaching and learning practices, the need to elaborate the internal dynamics of each mathematical process reflected in the learning outcome may affect the achievement of mathematical literacy. For this reason, as in Poland and other countries, it is recommended that the learning outcomes in the curricula should be arranged in a way that reflects a single mathematical literacy process. In cases where this arrangement is not made in the curricula, it can be said that it is a necessity to emphasise that its integration into the classroom is of vital importance. Although the outcomes and content of the curricula include mathematical literacy with all its processes in a holistic manner, it is the duty of teachers to effectively implement and transfer this curricula to students. In the literature, Höfer and Beckmann (2009), Altun and Akkaya (2014) and Lin and Tai (2015) state that one of the determining factors in the development of mathematical literacy is the role of mathematics teachers. Therefore, it should not be forgotten that teacher practices in the classroom have a great impact on student achievement as much as the renewal of the curricula. As a matter of fact, various studies emphasise the importance of teachers' mathematical literacy competencies in affecting their students' mathematics performance and draw attention to the importance of teachers' developing their own literacy skills (Botha, 2011; Genç, 2017).

This study, which was conducted to compare the distribution of the learning outcomes of the mathematics curricula of different countries in terms of mathematical literacy skills, was examined only depending on the learning outcomes specified in the curricula of the countries in line with the mathematical processes of mathematical literacy determined in 2022 PISA. For this reason, the study cannot fully understand how mathematics teaching in these countries is shaped in line with mathematical literacy. Based on the results obtained within the limits of the current study, it is suggested to create a holistic framework for developing mathematical literacy skills by examining the textbooks of the countries' mathematics lessons or making more detailed comparisons by observing the mathematics teaching in the learning environment. In addition, in line with the results obtained from this study, the suggestion of making necessary regulations in the curricula and presenting designs to support mathematical literacy skills will contribute to the literature.

**Compliance with Ethical Standards**

*Disclosure of potential conflicts of interest*

The authors declare that no processes involved in conducting this study have the potential for conflicts of interest.

*Funding*

No payments or scholarships have been received from any institution for this study.

*CRediT author statement*

First author is responsible for conceptualization methodology and writing, second author is responsible for reviewing and editing.

*Research involving Human Participants and/or Animals*

No data were obtained from human or animal subjects in this study. Data was collected only through document analysis. Therefore, ethical permission was not obtained.

## Farklı Ülkelerin Matematik Öğretim Programlarındaki Öğrenme Çıktılarının Niceliksel ve Matematik Okuryazarlığı Becerisi Bağlamında Niteliksel olarak Karşılaştırılması

**Özet:**

Farklı ülkelerin matematik öğretim programlarındaki öğrenme çıktılarının niceliksel ve niteliksel boyutta incelenmesi amacıyla nitel modelde yürütülen bu çalışmada, ülkelerin öğretim programlarına yönelik yazılı kaynaklar ve 2022 PISA sonuçları kullanılmıştır. Niceliksel boyutta; diğer ülkeler Türkiye'den daha az öğrenme çıktısıyla daha yüksek PISA puanı ve sıralamasına sahiptir. Niteliksel boyutta; Türkiye ve Polonya’da üçüncü, İsveç ve Danimarka’da ikinci matematik okuryazarlığı matematiksel sürecinde çıktıların yoğunlaştığı görülmüştür. Türkiye’deki çıktıların diğerlerine göre karmaşık yapıda olduğu ve matematiksel süreçlere dengeli dağılmadığı saptanmıştır. Bulgulara dayanarak matematik öğrenme çıktılarının azaltılması ve matematiksel süreçlere daha eşit dağılımının sağlanmasıyla yenilenerek bilginin yanında becerilerin de derinlemesine geliştirilebilmesine olanak tanıyan bir program önerilmektedir.

*Anahtar kelimeler:* Matematik okuryazarlığı becerisi, matematiksel süreçler, matematik öğretim programı, öğrenme çıktısı.

## References

Akpınar, Y. (2004). Teachers' approaches to some educational technology related factors affecting learning. *The Turkish Online Journal of Educational Technology*, *3*(3), 124–135. http://www.tojet.net/articles/v3i3/3315.pdf

Altun, M., & Akkaya, R. (2014). Mathematics teachers' comments on PISA math questions and our country's students' low achievement levels. *Hacettepe University Journal of Education*, *29*(1), 19-34. https://dergipark.org.tr/tr/download/article-file/87075

Altun, M. (2015). *Efemat 5-6* (1st ed.). Alfa Aktüel.

Ayvacı, H. Ş., & Durmuş, A. (2013). The problems faced by science and technology teachers in the first years of their profession and the change of these problems according to years. *The Journal of Atatürk University Kazım Karabekir Faculty of Education*, (27), 29-44. https://dergipark.org.tr/tr/pub/ataunikkefd/issue/2786/37413

Bal İncebacak, B. (2022). Comparison of the mathematics content of curriculum for primary school in Türkiye and Singapore. *Trakya Education Journal*, *12*(3), 1403-1425. https://doi.org/10.24315/tred.984222

Botha, J. J. (2011). *Exploring mathematical literacy: The relationship between teachers' knowledge and beliefs and their instructional practices.* University of Pretoria, South Africa. http://hdl.handle.net/2263/28984

Buluş Kırıkkaya, E. (2009). Opinions of science teachers in primary schools related to science and technology program. *Turkish Science Education Journal*, *6*(1), 133-148. https://www.tused.org/index.php/tused/article/view/113/70

Chung, I., & Chung, I. (2008). *Mathematics curriculum in pasific rim countries-Chine, Japan, Korea, and Singapore*. Research in mathematics education (pp.255-258). Information Age Publishing.

Çaycı, B. (2018). Evaluation of lesson duration and lesson hours of primary school according to opinions of class teachers. *International Journal of Eurasian Education and Culture*, *3*(5), 117-131. https://dergipark.org.tr/en/download/article-file/599551

Çetinbağ, A. (2019). *Comparing Turkey and Canada in the context of the program elements of primary school mathematics curriculum* (Publication No. 572471). [Master's thesis, Marmara University]. Council of Higher Education Thesis Center.

Çilingir, E., & Dinç Artut, P. (2016). Effect of realistic mathematics education approach on visual mathematics literacy perceptions and problem solving attitude of students.

*Turkish Journal of Computer and Mathematics Education (TURCOMAT)*, *7*(3), 578-600. https://doi.org/10.16949/turkbilmat.277872

Çilingir Altıner, E., & Dinç Artut, P. (2017). The effet of instruction based realistic mathematics education on elementary students' avhievement, visual mathematic literacy and problem solving attitude. *Marmara University Atatürk Faculty of Education Journal of Educational Sciences*, *46*(46), 1-19. https://doi.org/10.15285/maruaebd.279963

Çoban, A. (2011). *Comparison of primary mathematics curriculum of the United States of America, England and Turkey* (Publication No. 302987). [Master's thesis, Celal Bayar University]. Council of Higher Education Thesis Center.

Çobanoğlu, R., & Yıldırım, A. (2021). Curriculum development studies in turkey: a historical analysis from the declaration of republic to present. *The Journal of Turkish Educational Sciences*, *19*(2), 810-830. https://doi.org/10.37217/tebd.912329

Creswell, J. W. (2017) *Araştırma deseni: Nicel, nitel ve karma yöntem yaklaşımları*. (Demir, S. B. Trans.). Eğiten Kitap. https://www.egitenkitap.com/arastima-deseni-research-design-john-w.-craswell

Deveci, İ. (2018). Comparison of 2013 and 2018 science curricula in terms of basic elements in Turkey. *Mersin University Journal of the Faculty of Education*, *14*(2), 799-825.

https://doi. org/10.17860/mersinefd.342260

Dibek, M., Yalçın, S., & Yavuz, H. Ç. (2016). Investigation on the relationships between information communication technology and mathematics literacy for Turkey students. *The Journal of Ahi Evran University Kırşehir Faculty of Education*, *17*(3), 39-58. https://dergipark.org.tr/tr/download/article-file/1487412

Dursun, Ş., & Dede, Y. (2004). The factors affecting students' success in mathematics: mathematics teachers' perspectives. *The Journal of Gazi University Gazi Faculty of Education*, *24*(2), 217-230. https://dergipark.org.tr/tr/download/article-file/77327

Duygu, N. (2013). *International comparative study of elementary school mathematics curriculum* (Publication No. 351003). [Master's thesis, Gaziosmanpaşa University]. Council of Higher Education Thesis Center.

EMU. (2019). [Danmarks Læringsportal]. Matematik, Fælles Mål. https://www.emu.dk/grundskole/matematik/faghaefte-faelles-maal-laeseplan-og-vejledning?b=t5-t9

Erbilge, A. E. (2019). *Comparison of secondary school mathematics curriculum of Turkey, Canada and Hong Kong* (Publication No. 584838). [Master's thesis, Marmara University]. Council of Higher Education Thesis Center.

Erdoğan, F., Hamurcu, H., & Yeşiloğlu, A. (2016). The analyzing of TIMSS 2011 Turkey and Singapore results in mathematics program. *Cumhuriyet International Journal of Education-CIJE*, *5*(USOS Özel Sayı), 31-43. https://doi.org/10.30703/cije.321423

Genç, M. (2017). *Exploring secondary mathematics teachers' conceptions of mathematical literacy* (Publication No. 463568). [Master's thesis, Middle East Technical University]. Council of Higher Education Thesis Center.

Gürbüz, M. Ç. (2014). *The effect of pisa mathematical literacy teaching on writing Pisa questions and mathematical literacy* (Publication No. 372216). [Master's thesis, Uludağ University]. Council of Higher Education Thesis Center.

Güven, İ., & Gürdal, A. (2011). Comparative analysis of science education systems of Turkey and Canada. *Journal of Turkish Science Education*, *8*(4), 89-110. https://dergipark.org.tr/tr/download/article-file/1914

Güzel, İ., Karataş, İ., & Çetinkaya, B. (2010). A comparison of secondary school mathematics curriculum guidebooks: Turkey, Germany and Canada. *Turkish Journal of Computer and Mathematics Education*, *1*(3), 309-325. https://dergipark.org.tr/tr/download/article-file/201309

Hobson, A. (2001). Teaching relevant science for scientific literacy: adding cultural context to the sciences. *Journal of College Science Teaching*, *30*(4), 238-243. https://www.jstor.org/stable/42991213

Hook, W., Bishop, W., & Hook, J. (2007). A quality math curriculum in support of effective teaching for elementary schools. *Educational Studies in Mathematics*, *6*5(2), 125-148. https://doi.org/10.1007/s10649-006-9050-4

Höfer, T., & Beckmann, A. (2009). Supporting mathematical literacy: Examples from a cross-curricular project. *International Journal on Mathematics Education*, *41*(1), 223–230. http://doi.org/10.1007/s11858-008- 0117-9

Kahramanoğlu, R., Bay, E., Vural, Ö. F., & Aydın Aşk, Z. (2016). Teacher views on school-based curriculum development approach (an analysis of applicability in Turkey)**.** *Mustafa Kemal University Journal of Graduate School of Social Sciences*, *13*(34), 180-201. https://dergipark.org.tr/tr/download/article-file/226448

Karakaş, A., & Ezentaş, R. (2021). Planning, implementation and evaluation of mathematical literacy education provided to seventh grade students. The *Journal of National Education, 50*(232), 225-245. https://doi.org/10.37669/milliegitim.743329

Karataşlı, E. (2019). *A comparative study of Australia-Waldorf and Turkish secondary mathematics curriculum* (Publication No. 600572). [Master's thesis, Hacettep University]. Council of Higher Education Thesis Center.

Kaya, H. İ., & Karakaya, Ş. (2012). Effects of the practices based on constructivist learning in teacher education on teacher candidates' tendencies of problem solving. *Journal of Institute of Social Sciences*, 9, 79-95. https://www.kafkas.edu.tr/dosyalar/sobedergi/file/009/6.pdf

Köysüren, M., & Üzel, D. (2018). The effect of using technology in mathematics teaching to mathematical literacy of grade 6 students. *Necatibey Faculty of Education Electronic Journal of Science and Mathematics Education*, *12*(2), 81-101. https://doi.org/10.17522/balikesirnef.506418

Lidbäck, J. (2021). *Teachers' interpretations regarding the revisions coming with Lgr22*: *A study about the 2022 revision of the English syllabus, and the interpretations that English teachers have regarding the changes*. Karlstad University. https://www.diva-portal.org/smash/get/diva2:1565646/FULLTEXT01.pdf

Lin, S. W., & Tai, W. C. (2015). Latent class analysis of students′ mathematics learning strategies and the relationship between learning strategy and mathematical literacy. *Universal Journal of Educational Research*, *3*(6), 390–395. https://doi. org/10.13189/ujer.2015.030606

MoNE. (2019). *PISA 2018 Türkiye Ön Raporu.* Ministry of National Education, Ankara. https://www.meb.gov.tr/meb_iys_dosyalar/2019_12/03105347_pisa_2018_turkiye_on_raporu.pdf

MoNE. (2023). *PISA 2022 Türkiye Raporu.* Ministry of National Education, Ankara. https://pisa.meb.gov.tr/meb_iys_dosyalar/2024_03/21120745_26152640_pisa2022_rapor.pdf

Öçal, T. (2017). Comparing Turkish early childhood education curriculum with respect to common core state standards for mathematics. *Journal of Qualitative Research in Education*, *5*(3), 155-171. https://doi.org/10.14689/issn.2148- 2624.1.5c3s7m

Özgün-Koca, S. A., & Şen, A. İ. (2002). 3. Evaluation of the results of third international mathematics and science study for Turkey. *Hacettepe University Journal of Education, 23*,145-154. https://dergipark.org.tr/tr/download/article-file/87927

Podstawa (2017). Szkoła Podstawowa IV-VIII: Podstawa programowa kształcenia ogólnego dla szkoły podstawowej. https://podstawaprogramowa.pl/Szkola-podstawowa-IV-VIII

Podstawa (2018). Liceum/Technikum: Podstawa programowa kształcenia ogólnego dla czteroletniego liceum ogólnokształcącego i pięcioletniego technikum. https://podstawaprogramowa.pl/Liceum-techniku

Revina, S., & Leung, F. K. S. (2018). Educational borrowing and mathematics curriculum: Realistic mathematics education in the Dutch and Indonesian primary curriculum. *International Journal on Emerging Mathematics Education, 2*(1), 1-16. https://doi. org/10.12928/ijeme.v2i1.8025

Schoen, R., Erbilgin, E., & Hacıömeroğlu, S. E. (2011). Analyzing the next generation sunshine state standards for mathematics: Is the state curriculum still a mile wide and an inch deep? Dimensions in mathematics. *Spring*. *31*(1), 30-39. https://www.schoenresearch.com/wp-content/uploads/2020/08/schoen-ergilbin-haciomeroglu-2011-Analyzing-the-Next-Generation-Sunshine-State-Standards-for-mathematics-Is-the-state-curriculum-still-a-mile-wide-and-an-inch-deep.pdf

Skolverket (2022). Läroplan för grundskolan, förskoleklassen och fritidshemmet Lgr22. https://www.skolverket.se/publikationsserier/styrdokument/2024/laroplan-for-grundskolan-forskoleklassen-och-fritidshemmet---lgr22

Steinberg, R. N. (2011) *An inquiry into science education, where the rubber meets the road*. Sense. https://doi. org/10.1007/978-94-6091-690-8

Var, S. & Altun, M. (2021). Investigation of the learning process in the field measurement subject in the fourth grade mathematics course according to the constructivist learning theory. *Amasya Education Journal*, *10*(2), 1-23. https://doi. org/10.17539/amauefd.891703

Wojnak, J., & Majorek, M. (2018). Polish education system under 2017 reform: Assumptions, aims and controversies. SHS Web of Conferences ERPA 2018. https://doi. org/10.1051/shsconf/20184801043

Xie, X., & Carspecken, P. F. (2019). *Philosophy, learning and the mathematics curriculum: dialectal materialism and pragmatism related to Chinese and US mathematics curriculum*. Brill.

Yang, D. C., Tseng, Y. K., & Wang, T. L. (2017). A Comparison of geometry problems in middle-grade mathematics textbooks from Taiwan, Singapore, Finland, and the United States. *Eurasia Journal of Mathematics, Science and Technology Education*, *13*(7), 2841-2857. https://doi.org/10.12973/eurasia.2017.00721a

Yavuz Topaloğlu, M., & Balkan Kıyıcı, F. (2015), Comparison of science curriculum: Turkey and Australia. *Bartın University Journal of Faculty of Education*, *4*(2), 344- 363. https://doi. org/10.14686/buefad.v4i2.1082000266

Yazıcıoğlu, Ö., & Pektaş, M. (2018). A comparison of the middle school science programmes in Turkey, *Singapore and Kazakhstan. International Electronic Journal of Elementary Education*, *11*(2), 143-150. https:/doi.org/10.26822/iejee.2019248588

Yıldırım, A., & Şimşek, H. (2008). *Sosyal bilimlerde nitel araştırma yöntemleri*. Seçkin.

Necatibey Faculty of Education Electronic Journal of Science and Mathematics Education
Vol. 18, Issue 2, December 2024, pp. 351-374. ISSN: 1307-6086

Research Article

# The Effects of Formative Assessment on Student Achievement and Attitudes in Math*

**Çağla AKBAŞ[1], Ceyhun OZAN[2]**

[1] Ministry of National Education, TOBB Efkan Ala İmam Hatip Middle School, Türkiye, caglaakbas27@gmail.com, http://orcid.org/0000-0002-1474-1681

[2] Atatürk University, Kâzım Karabekir Faculty of Education, Türkiye, ozanceyhun@atauni.edu.tr, http://orcid.org/0000-0002-1415-7258

Received : 08.07.2024 Accepted : 22.10.2024

Doi: https://doi.org/10.17522/balikesirnef.1512149

---

*Abstract* – The objective of formative assessment is to allocate or assign a value to the student, identify the student's learning deficiencies and make instructional arrangements to eliminate these deficiencies. This study employed a non-equivalent group pre-test and post-test design to investigate the effects of formative assessment on fifth-grade students' academic achievement and their attitudes toward mathematics. The independent variables were formative assessment and level-determining evaluation methods. The dependent variables were students' academic achievements and attitudes toward mathematics. The research group comprised fifth-grade students of a determined secondary school in the Pasinler district of Erzurum province. Through random assignment, 17 students were included in the experimental group and 13 in the control group. The data were obtained using the achievement test and attitude scale. This research was conducted in a secondary school fifth-grade math class for 11 weeks. The mathematics teacher, the researcher, delivered lessons to both groups. The mathematics achievement of students in the experimental group was higher than the control group. However, the difference was not statistically significant. Similarly, no significant difference was found in the attitudes of students toward math. Various recommendations were made in light of the findings obtained in this study.

*Keywords*: Formative assessment, mathematics education, mathematics achievement, attitude to mathematics, elementary education, secondary school.

------------

Corresponding author: Ceyhun OZAN, ozanceyhun@atauni.edu.tr

* This study was produced from the master's thesis published by first author under the supervision of the second author. This paper was produced from the project supported by Atatürk University Scientific Research Projects Unit with the code SYL-2022-10227.

## Introduction

Formative assessment is a method that entails collecting and utilizing data on student learning to offer continuous feedback and assistance toward achieving academic objectives. Black and Wiliam (1998a) characterize formative assessment as “the approach employed by educators and learners to identify and address student learning to enrich that learning throughout the instructional process.” Put differently, formative assessment strives to enhance the teaching and learning experience by conducting evaluations and providing feedback that aligns with learning objectives throughout the learning journey. If a student's understanding is not checked, it becomes impossible to know what they have learned from the lesson (Kültür, 2021).

Formative assessment is an essential tool for teachers and students. It enables instructors to recognize struggling students early, foster student engagement and motivation, and promote more profound learning. Therefore, it should be an essential component of any curriculum. By providing feedback on their performance and guiding them to identify areas where they require improvement, formative assessment can promote profound learning, boost student motivation and engagement, and facilitate realizing their maximum potential (Black & Wiliam, 1998a; Hattie & Timperley, 2007; Sadler, 1989).

While formative assessment is recognized as a valuable tool for learners at all stages, fifth grade represents a crucial period in students' academic development. Students’ transition from elementary foundational learning to more complex cognitive tasks at this stage, making it an ideal time to examine how formative assessment can support this critical phase. Research has shown that formative assessment is especially effective in promoting student engagement and improving outcomes during this developmental stage, where the demands of the curriculum increase and students are expected to demonstrate higher-order thinking skills (Black & Wiliam, 1998a). Therefore, focusing on fifth-grade students allows the study to address this critical juncture in learning, providing valuable insights into how formative assessment can facilitate academic achievement during a pivotal time in students' educational trajectories.

### Features of Formative Assessment

As noted by Black et al. (2003), formative assessment comprises four fundamental elements: (i) Clarifying learning objectives and achievement criteria, (ii) Promoting the quality of questioning/dialogue, (iii) Improving the quality of scoring/feedback/record keeping (iv) Using peer and self-assessment. The "big idea" is to leverage insights into student learning to adapt instruction for enhanced responsiveness to individual student needs. In other words, instruction is adaptable to the student's learning needs.

Clarifying learning objectives and achievement criteria is fundamental to effective teaching and learning practices. Learning objectives refer to the goals students should achieve at the end of a lesson or unit, while criteria for achievement refer to the standards or expectations for achieving those objectives (Black & Wiliam, 1998a; Rodriguez & Albano, 2017). Hattie (2009) contends that the effective communication and clarification of learning objectives and achievement criteria for students represent a robust pedagogical strategy for improving student achievement. More recent studies, such as those by Brookhart (2017), have expanded on this by emphasizing that clear criteria improve academic outcomes and foster student autonomy.

Questioning and dialogue are essential components of formative assessment (Hodgson & Pyle, 2010; Walsh, 2022). Borich (2014) suggests that in a typical primary or secondary school class hour, 50 or more questions may be asked and that up to 80 percent of classroom time can be devoted to questioning and answering. Through meaningful dialogues, educators and learners can engage in an interactive exchange of ideas to resolve misunderstandings and promote a more profound level of learning (Black & Wiliam, 1998a; Heritage, 2007). Walsh and Sattes (2016) highlight that high-quality questioning promotes engagement and fosters metacognition, as students reflect on their thinking processes.

For formative assessment to be practical, improving the quality of scoring, feedback, and record-keeping is imperative. As Hattie and Timperley (2007) emphasize, feedback is a central formative assessment component. Andrade et al. (2015) contend that the efficacy of formative assessment depends on the quality of feedback given to students regarding their learning progress and teachers regarding their instructional practices. Feedback on formative assessment wields a significant influence on both learning and achievement. Some researchers even state that feedback is the most crucial element of students' learning (Bell, 2007; Brown, 2018; Shavelson et al., 2008; Wiliam, 2018).

Self and peer assessment are other effective formative assessment components (Nicol & Macfarlane-Dick, 2006). The responsibility for offering feedback is not exclusive to teachers. When properly instructed and guided, self and peer assessments can also provide students with valuable feedback focused on learning (Andrade, 2019; Andrade et al., 2015; Huisman et al., 2019; Panadero et al., 2018). Peer assessment or feedback involves students providing comments on the work of their peers (Topping, 2009).

**Formative Assessment in Mathematics Education**

The literature highlights that formative assessment is a powerful tool in mathematics education, as it facilitates active engagement and provides real-time feedback, both of which are essential for learning (Black & Wiliam, 1998b; Hattie & Timperley, 2007). However, a deeper exploration into what has been written specifically about formative assessment within the context of mathematics education is necessary. Research emphasizes that mathematics plays a pivotal role in developing critical problem-solving abilities and logical reasoning skills, which are crucial for academic and real-world success (Boaler, 2016; Schoenfeld, 2013). Given the abstract nature of many mathematical concepts, formative assessment provides a scaffold for students to make connections between different concepts and correct their misunderstandings before they become entrenched (Swan, 2006).

Despite its significance, many students may face significant challenges with mathematics, leading to poor academic performance, decreased self-efficacy, and lower motivation (Pajares & Graham, 1999; Wigfield & Eccles, 2000). Therefore, by providing regular feedback and helping teachers adjust instruction based on students' progress, formative assessment has been proven to support learning in this field (Black & Wiliam, 1998a; Wiliam, 2011).

In mathematics, formative assessment practices include a variety of strategies such as diagnostic questions, exit tickets, and real-time quizzes, which allow teachers to gather data about student progress. These tools help identify common misconceptions (Fennell, 2011) and provide an opportunity to differentiate instruction, ensuring that all students receive the support they need to succeed (Clarke, 2005). However, this study needs to further engage with what previous studies have specifically revealed about formative assessment in mathematics to underscore its contribution to the existing body of knowledge.

Studying formative assessment in mathematics is crucial because it addresses academic and affective challenges students face in this subject. The cyclical process of data collection, analysis, and instructional adjustment inherent in formative assessment (Black & Wiliam, 1998a) is particularly beneficial for mathematics, where frequent misconceptions can hinder students' understanding. Recognizing these misconceptions and addressing them promptly can significantly improve learning outcomes. Since mathematics is a cumulative discipline, early identification of gaps through formative assessment can prevent future difficulties and enhance students' problem-solving and logical reasoning abilities. The study's focus on the practical application of formative assessment techniques such as problem exploration, group assessments, and peer collaboration (Lee, 2006) shows how tailored feedback and instructional adjustments can lead to better outcomes in mathematics education. These techniques highlight formative assessment's potential to foster a deeper understanding of mathematics and its relevance to ongoing educational reforms.

While there is substantial research supporting the benefits of formative assessment for student achievement in mathematics (Black & Wiliam, 1998a; Hattie & Timperley, 2007; Heritage et al., 2009; Sadler & Good, 2006), this study adds value by addressing specific strategies and tools that can be used to identify students' learning gaps in real-time. Furthermore, by focusing on formative assessment's impact on students' self-efficacy (Schiefele et al., 2012), problem-solving skills (Cavanagh et al., 2016; Kramarski & Michalsky, 2009; Beatty & Gerace, 2009), and motivation (Hounsell et al., 2008), this study extends understanding of how formative assessment can influence not only academic outcomes but also students' emotional and cognitive engagement with mathematics. The present study emphasizes that formative assessment is not just about improving test scores but about fostering a growth mindset and creating a more inclusive learning environment that supports diverse learners. The findings of this research will help bridge the gap between theory and practice, offering educators practical tools to enhance students' learning experiences in mathematics classrooms.

## Research Questions

This research investigates the influence of fifth-grade students' formative assessment on mathematics achievements and attitudes toward mathematics in secondary schools. To address this objective, this study outlines the following research questions:

1) Is there a significant difference between the control group's pre-test and post-test mathematics achievements?
2) Is there a significant difference between the experimental group's pre-test and post-test mathematics achievements?
3) Is there a significant difference between the control group's pre-test and post-test attitudes toward math?
4) Is there a significant difference between the experimental group's pre-test and post-test attitudes toward math?

5) Is there a significant difference between the experimental and control groups' post-tests?

## Method

### Research Design

In this research, we utilized a non-equivalent group pre-test and post-test design. This design is classified as experimental because it allows for comparing outcomes between an experimental group that receives the intervention (formative assessment) and a control group that does not. According to Campbell et al. (1963), pre-test and post-test measures help establish a baseline for assessing the effectiveness of the intervention while controlling for pre-existing differences between groups.

Due to logistical constraints, it was not feasible to randomly assign individuals to either the experimental or control groups. However, we made every effort to ensure that participants were assigned to the groups as randomly as possible, thereby minimizing potential biases and confounding factors (Shadish et al., 2002). This approach aligns with the principles outlined by Creswell (2014), emphasizing the importance of quasi-experimental designs when random assignment is impractical, yet the need for causal inferences remains.

### Study Group

The study group consisted of fifth-grade students enrolled in a secondary school located in the Pasinler district of Erzurum province, Türkiye. The selection of the study group was based on the researcher's accessibility and the possibility of attending mathematics classes and extracurricular activities with the group on an individual basis. The researcher's presence throughout the application process enabled the observation of students' interests, attitudes, and behaviors. Although the study group's district is not particularly large, it can be described as having a low-to-middle income level in general. Most students in the study group have a middle-income level, while their parents' education levels are mainly at the primary school level. The average class size throughout the county is currently 16. A total of 30 students participated in this study, with 17 from the 5/B class and 13 from the 5/C class in the school representing fifth-grade students in the county. Demographics of the study group are shown in Table 1.

**Table 1** Demographics of Study Group

| Gender | Control | Experimental | Total |
|---|---|---|---|
| Female | 5 | 10 | 15 |
| Male | 8 | 7 | 15 |
| Total | 13 | 17 | 30 |

The results regarding whether there was a significant difference between the mathematics achievement and attitude scores of the experimental and control groups before the experimental procedures are shown in Table 2.

**Table 2** Results of Pre-tests

| Group | n | $\bar{X}$ | SS | sd | t | p |
|---|---|---|---|---|---|---|
| Control (achievement) | 13 | 4.61 | 1.98 | 28 | 1.830 | .078 |
| Experimental (achievement) | 17 | 5.88 | 1.79 | | | |
| Control (attitude) | 13 | 3.78 | .37 | 28 | 2.798 | .009 |
| Experimental (attitude) | 17 | 4.22 | .49 | | | |

As shown in Table 2, it was determined that there was no significant difference between the mathematics achievement scores of the experimental and control group students ($t(28) = 1.830$; $p > .05$). A significant difference was found in the mathematics attitude scores between the experimental and control groups ($t = 2.798$; $p < .05$). The attitude scores of the experimental group were significantly higher before the experimental procedures.

**Process**

This research was conducted over the course of 11 weeks in a fifth-grade mathematics class at a secondary school. The lessons were taught by the mathematics teacher, the researcher, to both groups. It should be noted that formative assessment practices were only applied in the experimental group, while no such practices were implemented in the control group. The lessons in the experimental group were designed with the intention of achieving the same learning outcomes, while also incorporating formative assessment applications. The teacher who conducted the practices in this study is a graduate student. Before the experimental procedures began, the teacher received weekly guidance from the thesis advisor on formative assessment practices. This training lasted one hour per week throughout the entire semester, ensuring that the teacher had a clear understanding of how to implement formative assessment techniques during this study effectively. However, it is recommended that more comprehensive and structured professional development programs be provided to teachers before such studies to enhance their capacity to apply formative assessment effectively.

Black et al. (2003) identified four key elements of formative assessment: (i) the clarification of learning objectives and criteria for achievement, (ii) the promotion of the quality of questioning and dialogue, (iii) the improvement of the quality of scoring and feedback, and (iv) the utilisation of peer and self-assessment. The teaching process of Ozan and Kıncal's (2018) research included the following practices within the framework of the four fundamental components of formative assessment:

***The Clarification of Learning Objectives and Criteria for Achievement***

This study employed the following strategies to elucidate the learning objectives and achievement criteria: (i) The teacher introduced the learning objectives of each lesson and engaged in a discussion with the students about the expected learning outcomes. (ii) The teacher frequently reminded the students of the learning objectives throughout the lesson. (iii) After completing the lesson, the teacher discussed with the students what they had learned and how it related to the learning objectives. (iv) The criteria for achievement in the activities that students would undertake in the classroom were made clear, including the necessary actions and the means of attaining them. (v) The criteria for completing homework assignments successfully were also made known to the students.

***The Promotion of the Quality of Questioning and Dialogue***

This study employed the following strategies to enhance the quality of questioning and dialogue: (i) Collaborative group work was employed to facilitate student dialogue. (ii) The teacher formed groups heterogeneously, considering factors, such as gender, academic performance, and affective traits of the students. The study groups consisted of four groups with four students each and one group with five students. (iii) At the conclusion of each unit, the students were assigned to new groups through random selection. Students were encouraged to provide feedback to each other during group work. (iv) The teacher employed high-level thinking skill questions throughout the course. (v) Students were allotted time to reflect before answering questions, with the amount of reflection time varying from 3 to 25 seconds, based on the complexity of the question. (vi) As part of assessing their higher-level thinking abilities, students were given the opportunity to engage in peer discussions. (vii) The "No Hands Up" strategy was employed to the teacher's questions. This strategy aimed to enhance the participation and engagement of all students by prohibiting them from raising their hands, thereby requiring each student to contemplate the answer and participate in the lesson.

### *The Improvement of the Quality of Scoring and Feedback*

This study employed the following strategies to enhance the quality of scoring, feedback: (i) Instead of assigning grades or points, the teacher primarily provided students with immediate feedback on their classwork or homework. (ii) Students were given the opportunity to amend their work in accordance with the feedback provided by the teacher. (iii) Quizzes were administered at the conclusion of each topic and unit to identify students' learning gaps and provide feedback. (iv) Students were provided with feedback on their weaknesses following the quiz. (v) Based on the feedback students received on their quizzes, assignments, and activities, adjustments were made in the classroom. Many arrangements were made to address the identified weaknesses, including small group work, individual activities, re-explaining, worksheets, teaching by showing and doing, internet research, clarification of assessment criteria, additional reading, and concept mapping. Efforts were made to create a learning-oriented classroom culture with student-centered practices in which students are active. (vi) Both summative and formative assessment data were recorded when assessing students. Student names were converted into codes. (vii) The assessment data on the students' performance were shared with the following year's teacher to ensure continuity in their learning. (viii) Rewards were given to recognize and encourage students' progress. (ix) The awards were not used to reward students by comparing them with each other but by considering their development levels. Rewards were provided to students who demonstrated individual progress in quizzes.

### *The Utilization of Peer and Self-assessment*

This study incorporated the following methods for utilizing self and peer assessment: (i) The regular integration of self and peer assessments into the lessons. Self-assessment and peer assessment activities were carried out after each unit. (ii) Guidelines were presented for each activity, and students were educated on performing self and peer assessments. (iii) The completed assignments were discussed among the students to assess whether they met the expected standards.

The control group’s instructional sessions followed the standard curriculum without incorporating formative assessment strategies. As a result, several key elements typically associated with formative assessment were absent. For example, no cooperative group work was facilitated, and high-level thinking questions were rarely posed. Additionally, there were no short quizzes to gauge student understanding regularly, no awards or incentives were provided to motivate student performance, and self-assessment and peer assessment activities were not part of the instructional process. This traditional approach focused on delivering the material as per the curriculum guidelines without the additional scaffolding formative assessment provides.

## Data Collection Tools

### *Mathematics Achievement Test*

The Mathematics Achievement Test was used to measure the mathematical achievement of students in the experimental and control groups. The test was developed based on the learning outcomes delineated in the "Mathematics Curriculum," published by the Republic of Türkiye Ministry of National Education (MoNE, 2017). The test encompassed 37 questions about 17 learning outcomes. The opinions of three mathematics educators from the same educational institution and a curriculum and instruction expert were sought to ensure the test's reliability. Following careful consideration, the experts concluded that no changes were required. The test was administered to sixth-grade students who had received mathematics education the previous year. Item analysis was performed after a preliminary application of the test to 22 sixth-grade students, and item difficulty and discrimination levels were determined. We created 27 percent lower and upper groups to calculate item discrimination. Table 3 shows the item difficulty and discrimination indexes of the pre-application test.

**Table 3** Pre-application of the Achievement Test's Degree of Item Difficulty and Discrimination

| Item number | Difficulty index | Discrimination index | p | Item number | Difficulty index | Discrimination index | p |
|---|---|---|---|---|---|---|---|
| 1 | 0.90 | -0.16 | .25 | 20 | 0.13 | -0.16 | .35 |
| 2* | 0.63 | 0.66 | .00 | 21* | 0.68 | 0.66 | .00 |
| 3 | 0.50 | 0.33 | .00 | 22 | 0.54 | 0.33 | .00 |
| 4* | 0.77 | 0.5 | .00 | 23 | 0.31 | 0.16 | .28 |
| 5 | 0.22 | 0.33 | .00 | 24* | 0.50 | 0.33 | .00 |
| 6* | 0.31 | 0.33 | .00 | 25 | 0.45 | 0.33 | .00 |
| 7* | 0.27 | 0.33 | .00 | 26* | 0.09 | 0.33 | .00 |
| 8 | 0.04 | 0.16 | .30 | 27 | 0.18 | -0.16 | .30 |
| 9* | 0.18 | 0.50 | .00 | 28 | 0.27 | 0 | .40 |
| 10 | 0.50 | 0.33 | .00 | 29* | 0.59 | 0.50 | .00 |
| 11* | 0.63 | 0.33 | .00 | 30 | 0.40 | 0.16 | .20 |
| 12 | 0.22 | 0.16 | .27 | 31 | 0.54 | 0.16 | .15 |
| 13* | 0.27 | 0.66 | .00 | 32 | 0.63 | 0 | .50 |
| 14 | 0.22 | 0 | .40 | 33* | 0.50 | 0.50 | .00 |
| 15* | 0.68 | 0.83 | .00 | 34* | 0.72 | 0.50 | .00 |
| 16 | 0.54 | 0.50 | .00 | 35 | 0.77 | 0.50 | .00 |
| 17 | 0.77 | 0.33 | .00 | 36 | 0.81 | 0.50 | .00 |
| 18* | 0.22 | 0.50 | .00 | 37* | 0.72 | 0.66 | .00 |
| 19* | 0.5 | 0.50 | .00 | | | | |

*Items taken to the final test

According to Table 3, the degree of item discrimination varied from 0.16 to 0.83. Items with discrimination of less than 0.30 were excluded from the test. When selecting one question from each learning outcome based on their high degrees of discrimination, the discrimination values of the 17 selected items ranged from 0.33 to 0.83. The mean discrimination of the 17 selected items was 0.50. Each item in the test was assigned a point value, and the total score ranged from 0 to 17. The reliability of the achievement test was assessed by calculating the KR-20 reliability coefficient, yielding a value of 0.71. Tests with a KR-20 coefficient greater than 0.7 were considered reliable. Therefore, the achievement test is reliable.

### *Attitude Scale toward Mathematics*

The assessment of shifts in students' attitudes toward mathematics before and after the intervention relied on the utilization of “The Attitude Scale towards Mathematics” (ASTM), formulated by Önal (2013). This scale comprises 22 items and four factors: anxiety, interest, necessity, and work. There are 11 negative items on ASTM. The scale items are presented utilizing a 5-point Likert scale format, where respondents can choose from options such as “strongly disagree,” “disagree,” “undecided,” “agree” and “strongly agree.”

The overall scale demonstrated a Cronbach’s alpha of 0.90. The Cronbach's alpha coefficients of the dimensions of anxiety (5 items), interest (10 items), necessity (3 items), and working (4 items) were 0.74, 0.89, 0.70, and 0.69, respectively (Önal, 2013). In this research, Cronbach’s alpha for the overall scale was 0.82.

## Data Analysis

The normality of the data collected from the control and experimental groups was assessed using the Shapiro-Wilk normality test. The Shapiro-Wilk test is known for its high sensitivity in detecting deviations from normality, making it particularly effective for small sample sizes (Razali & Wah, 2011). Table 4 shows the Shapiro-Wilk results.

**Table 4** Shapiro-Wilk Results

| Test | Group | Statistic | p |
|---|---|---|---|
| Mathematics achievement pre-test | Control | 0.960 | .75 |
| | Experimental | 0.939 | .31 |
| Mathematics achievement post-test | Control | 0.894 | .11 |
| | Experimental | 0.954 | .52 |
| Mathematics attitude pre-test | Control | 0.949 | .58 |
| | Experimental | 0.910 | .10 |
| Mathematics attitude post-test | Control | 0.922 | .26 |
| | Experimental | 0.960 | .63 |

The Shapiro-Wilk test results indicated that both the control and experimental groups exhibited normality in their mathematics achievement and attitude scores for both pre-tests and post-tests, as all p-values were more significant than 0.05. After confirming that the collected data showed normal distribution and had equal variance, we used independent samples t-test to compare the mean differences between the control and experimental groups in terms of mathematics achievement and attitudes toward mathematics and analysis of covariance (ANCOVA) to compare the mean differences between the control and experimental groups post-tests. ANCOVA helps assess the effect of the treatment while accounting for the initial differences in pre-test scores, thereby providing a clearer understanding of the impact of formative assessment on the students' achievements and attitudes (Tabachnick & Fidell, 2019).

**Threats to Internal Validity in Experimental Designs**

In this study, the authors acknowledge that the research design may be considered weak due to the use of non-equivalent groups. Non-equivalent groups can introduce threats to internal validity, including selection bias, maturation effects, and history effects. To address these potential threats and enhance the robustness of the findings, the authors implemented several strategies:

1. Selection Bias: Although random assignment to groups was not feasible due to logistical constraints, efforts were made to ensure that participants were assigned to groups as randomly as possible within the given classes. This minimizes the likelihood that pre-existing differences between the groups would confound the results.
2. Maturation Effects: To control potential maturation effects, the study was conducted over a relatively short duration (11 weeks). Additionally, pre-test scores were used to assess the initial equivalence of groups, allowing for a more accurate comparison of post-test outcomes.
3. History Effects: The authors took measures to reduce the impact of external events that might influence student learning. By conducting the study in a controlled classroom environment and maintaining consistent teaching practices across both groups, the authors aimed to limit the influence of outside factors on the students' achievement and attitudes toward mathematics.

4. Regular Monitoring and Feedback: Throughout this study, regular assessments were conducted to monitor student progress and provide feedback. This ongoing evaluation helped identify any unforeseen issues affecting internal validity, allowing for timely interventions.

## Findings

Table 5 shows descriptive statistics regarding the scores of both groups, along with the pre-test and post-test scores derived from the Achievement Test and ASTM.

**Table 5** Descriptive Statistics

| | Groups | n | Mean | Adjusted mean |
|---|---|---|---|---|
| Achievement (Pre-test) | Control | 13 | 4.61 | |
| | Experimental | 17 | 5.88 | |
| ASTM (Pre-test) | Control | 13 | 3.78 | |
| | Experimental | 17 | 4.22 | |
| Achievement (Post-test) | Control | 13 | 7.61 | 7.69 |
| | Experimental | 17 | 9.58 | 9.01 |
| ASTM (Post-test) | Control | 13 | 3.75 | 3.99 |
| | Experimental | 17 | 4.09 | 4.00 |

As indicated in Table 5, the control group students exhibited a pre-test achievement score of 4.61, which showed an increase to 7.61 in the post-test. In contrast, the experimental group students started with a pre-test achievement score of 5.88, demonstrating an improvement to 9.58 in the post-test. In terms of the ASTM scores, the control group students began with a pre-test score of 3.78, rising to 7.61 in the post-test.

There was a significant difference in the mathematics achievement scores among the control group students in favor of the post-test ($t(12) = -2.98$; $p < .05$). There was no significant difference in the attitude scores toward mathematics among the control group students ($t(12) = 2.67$; $p > .05$).

There was a significant difference in the mathematics achievement scores among the experimental group students in favor of the post-test ($t(16) = -5.370$; $p < .05$). There was no significant difference in the attitude scores toward mathematics among the experimental group students ($t(16) = 1.284$; $p > .05$). The statistical significance of the difference in students' post-test achievement scores was investigated using ANCOVA, with the results detailed in Table 6.

**Table 6** Post-test Scores of Achievements

| Source | Sum of squares | df | Mean square | F | p |
|---|---|---|---|---|---|
| Corrected Model | 85.920 | 3 | 28.640 | 2.979 | .050 |
| Intercept | 88.559 | 1 | 88.559 | 9.212 | .005 |
| Pre-test | 32.858 | 1 | 32.858 | 3.418 | .076 |
| Group | 10.810 | 1 | 10.810 | 1.125 | .299 |
| Error | 249.946 | 26 | 9.613 | | |
| Total | 2624.000 | 30 | | | |
| Corrected total | 335.867 | 29 | | | |

Table 6 shows no significant difference between the post-test achievement scores of the groups ($F_{(1, 26)} = 1.125; p > 0.05$). As a result, the implemented procedures did not yield a statistically significant difference in the achievement of the experimental group students compared to the control group students. Table 7 shows the ANCOVA results for ASTM post-test scores.

**Table 7** Post-test Scores of Attitudes

| Source | Sum of squares | df | Mean square | F | p |
|---|---|---|---|---|---|
| Corrected model | 4.176 | 3 | 1.392 | 8.773 | 000 |
| Intercept | .305 | 1 | .305 | 1.925 | .177 |
| Pre-test | 2.757 | 1 | 2.757 | 17.377 | .000 |
| Group | .271 | 1 | .271 | 1.705 | .203 |
| Error | 4.125 | 26 | .159 | | |
| Total | 476.017 | 30 | | | |
| Corrected total | 8.301 | 29 | | | |

There was no significant difference in the post-test scores regarding the attitude toward mathematics between the groups ($F_{(1, 26)} = 1.705, p > 0.05$).

## Conclusion, Discussion and Recommendations

In the present study, the observed difference was not statistically significant despite the experimental group, where formative assessment interventions were implemented, demonstrating higher mathematics achievement than the control group. In the literature, several studies have shown that formative assessment and feedback positively affect students' achievement. These studies include systematic reviews or meta-analyses studies, such as those by Black and Wiliam (1998a), Hattie and Timperley (2007), Shute (2008), and Kluger and DeNisi (1996). Moreover, recent studies have also focused on the effects of formative assessment on specific skills and subjects (Bulunuz et al., 2017; Gedikli, 2018; Kıncal & Ozan, 2018; Kültür, 2021; Ozan & Kıncal, 2018; Sönmez, 2020; Unaş, 2021).

In this study, the authors recognize that the absence of a significant difference in student achievement and attitudes toward mathematics may be attributed to several factors. First, the literature reveals that some studies indicate no significant effect of formative assessment on student achievement (Andrews, 2011; Collins, 2012; King, 2003; Tuominen, 2008; Yin et al., 2008). This suggests that the effectiveness of formative assessment can vary depending on various contextual and instructional factors.

One major challenge in the effective implementation of formative assessment is its integration into classroom instruction. Successful enactment requires that teachers understand effective formative assessment practices and have access to the necessary resources and support. The authors emphasize that teachers must receive adequate training to utilize formative assessment effectively (Black & Wiliam, 1998b; Earl, 2003; Heritage, 2007; Popham, 2008; Ruiz-Primo et al., 2008).

Furthermore, implementing formative assessment can demand additional time and effort from teachers, which may not always be feasible within the existing curriculum constraints. For example, the authors noted that for students with lower overall achievement to benefit from formative assessment fully, the time allocated to subjects might need to be increased.

In addition, the lack of access to essential technologies, such as smart boards and internet resources in classrooms, may further hinder the development of effective formative assessment practices. Without these tools, teachers may struggle to create engaging and interactive learning environments conducive to formative assessment, which could explain the lack of significant differences observed in the study's findings. By highlighting these challenges, the authors underscore the complexity of implementing formative assessment in educational settings and suggest that further research is needed to explore these dynamics.

Despite incorporating formative assessment practices within the experimental group in this study, no statistically significant difference was observed in the mathematics achievement of students when compared to the control group. Similarly, no significant difference was noted in the attitudes of students in the experimental group toward math than those in the control group. It is crucial to recognize that the pivotal components of formative assessment encompass feedback and self-assessment. In line with the feedback given to students within the scope of formative assessment, students must make various arrangements and try to guide themselves in line with their self-assessment. Students need time to develop these skills. Achievement did not increase at a statistically significant level in this study because the students needed to follow the regulations adequately after the feedback given by the teacher or the duration of one semester required to be increased. Koçak (2021) conducted a study with geography teachers and found that while teachers' attitudes toward formative assessment were positive, there were deficiencies in theoretical and practical terms, and students could not be sufficiently motivated to this assessment without grade concern. The existing education system's focus on result-oriented assessment approaches was also stated as a difficulty in practice.

The math curriculum issued by the Ministry of National Education in Türkiye (MoNE, 2017) includes only cognitive gains, and it is challenging to achieve affective gains for math in a curriculum that does not give place to affective gains. Students in Türkiye mostly think that mathematics is a complex subject, which creates anxiety and a negative attitude toward the subject. Baykul (2005) suggests that changing students’ attitudes towards math is not straightforward and is a long-term process. Therefore, implementing formative assessment practices over an extended period and not only for certain subjects but also for all the subjects in the curriculum with a formative approach may be more effective in changing students' attitudes toward the lesson.

The main shortcomings of formative assessment are time and resource limitations (Black & Wiliam, 1998a; Brookhart, 2013; Heritage, 2010; McMillan & Hearn, 2008; Popham, 2011; Sadler, 1998). Formative assessment can be dysfunctional when teachers cannot devote sufficient time to student feedback. In this case, teachers may be unable to adequately assess students’ performance or provide feedback. Another notable limitation is the teachers’ insufficient knowledge and experience regarding formative assessment methods. In this case, teachers may not assess students correctly or give incorrect feedback (Chappuis, 2009).

This study highlights the critical role of formative assessment in enhancing students' learning outcomes and improving teaching practices. Formative assessment has the potential to elevate student performance by providing real-time feedback and insights into their understanding. However, this study found no statistically significant differences between groups, suggesting that the effectiveness of formative assessment may depend on various contextual factors.

The absence of statistically significant differences in this study may stem from several factors. First, the limited training and teachers’ knowledge in implementing effective formative assessment strategies can hinder its impact. Additionally, logistical constraints, such as insufficient time for providing feedback and inadequate resources like access to technology, may have affected the implementation process. Moreover, the small sample size and the non-equivalent nature of the groups could also limit the generalizability of the findings.

Several limitations may influence the study's results. The non-equivalent group design poses a threat to internal validity, as it does not control for pre-existing differences between the experimental and control groups. The reliance on self-reported data for attitudes toward math may also introduce bias. Furthermore, the researcher’s dual role as the teacher may lead to potential biases in assessment and feedback. Finally, external factors, such as classroom dynamics and students' home environments, were not controlled for, which could have impacted the outcomes.

For further research, it is recommended that studies explore the long-term effects of formative assessment on student achievement and attitudes across diverse educational contexts. Additionally, investigating the impact of teacher training programs on the effective implementation of formative assessment could provide valuable insights. Further studies should also consider using a larger sample size and random assignment to strengthen the validity of the findings. Finally, examining the role of technology in facilitating formative assessment practices could be beneficial, particularly in addressing the challenges of accessibility and engagement.

In conclusion, while formative assessment is a powerful tool for enhancing learning and teaching, its effectiveness is contingent upon proper implementation, adequate training, and supportive resources. By addressing the identified limitations and building on the study's findings, educators and researchers can better harness the potential of formative assessment in educational settings.

**Compliance with Ethical Standards**

*Disclosure of potential conflicts of interest*

No conflict of interest.

*Funding*

None.

*CRediT author statement*

The article was collaboratively written by two authors, with each contributing equally to its content.

*Research involving Human Participants and/or Animals*

The ethics committee permissions were obtained with the letter dated 11.11.2021 and numbered 12/10 of the “Social and Human Sciences Scientific Research and Publication Ethics Committee of University Institute of Educational Sciences.”

## Matematik Eğitiminde Biçimlendirici Değerlendirmenin Öğrencilerin Akademik Başarı ve Tutumlarına Etkisi

**Özet:**

Biçimlendirici değerlendirmenin amacı öğrenciye bir değer biçmek değil, öğrencinin öğrenme eksikliklerini belirlemek ve bu eksiklikleri gidermeye yönelik öğretimsel düzenlemeler yapmaktır. Araştırmanın temel amacı, matematik öğretimi sürecinde biçimlendirici değerlendirmenin beşinci sınıf öğrencilerinin akademik başarıları ve matematik dersine yönelik tutumlarına etkisini incelemektir. Bağımsız değişken, biçimlendirici değerlendirme ile düzey belirleyici değerlendirme uygulamalarıdır. Bağımlı değişkenler ise öğrencilerin akademik başarıları ve matematiğe karşı tutumlarıdır. Araştırmanın çalışma grubunu Erzurum ili Pasinler ilçesinde belirlenen bir ortaokulun beşinci sınıf öğrencileri oluşturmaktadır. Yansız atama yoluyla deney grubunda 17, kontrol grubunda 13 öğrenci yer almıştır. Verilerin elde edilmesinde başarı testi ve tutum ölçeği kullanılmıştır. Çalışma, ortaokul beşinci sınıf matematik dersinde 11 hafta boyunca yürütülmüştür. Dersler hem kontrol hem de deney gruplarında aynı zamanda araştırmacı olan matematik öğretmeni tarafından verilmiştir. Araştırma sonuçlarına göre, biçimlendirici değerlendirme uygulamalarının olduğu deney grubundaki öğrencilerin matematik dersi başarıları, kontrol grubundaki öğrencilere göre yüksek bulunmuştur ancak elde edilen fark istatistiksel olarak anlamlı değildir. Benzer şekilde, kontrol ve deney grubunun matematik dersi tutumlarında da anlamlı bir fark bulunmamıştır. Araştırma sonucunda çeşitli önerilerde bulunulmuştur.

*Anahtar kelimeler*: Biçimlendirici değerlendirme, matematik eğitimi, matematik başarısı, matematik tutumu, ortaokul eğitimi, ortaokul.

**References**

Andrade, H., & Lui, A., Palma, M., & Hefferen, J. (2015). Formative assessment in dance education. *Journal of Dance Education, 15(2)*, 47-59. https://doi.org/10.1080/15290824.2015.1004408

Andrade, H. L. (2019). A critical review of research on student self-assessment. *Frontiers in Education*, *4*, 1-13. https://doi.org/10.3389/feduc.2019.00087

Andrews, T. L. (2011). *The use of goal setting and progress self-monitoring with formative assessment in community college to increase academic achievement and self-efficacy* [Unpublished doctoral dissertation]. Temple University.

Baykul, Y. (2005). *Teaching mathematics in primary education (for grades 1–5)*. Pegem Academy.

Beatty, I. D., & Gerace, W. J. (2009). Technology-enhanced formative assessment: A research-based pedagogy for teaching science with classroom response technology. *Journal of Science Education & Technology*, *18*(2), 146. http://doi.org/10.1007%2Fs10956-0089140-4

Bell, B. (2007). Classroom assessment of science learning. In S. Abell, & N. Lederman (Eds.), *Handbook of research on science education* (pp. 965–1006). Lawrence Erlbaum.

Black, P., Harrison, C., Lee, C., Marshall, B., & Wiliam, D. (2003). *Assessment for learning: Putting it into practice*. Open University.

Black, P., & Wiliam, D. (1998a). Assessment and classroom learning. *Assessment in Education*, *5*, 7-74. https://doi.org/10.1080/0969595980050102

Black, P., & Wiliam, D. (1998b). Inside the black box: Raising standards through classroom assessment. *Phi Delta Kappan*, *80*(2), 139-144. https://kappanonline.org/inside-the-black-box-raising-standards-through-classroom-assessment/

Boaler, J. (2016). *Mathematical mindsets: Unleashing students' potential through creative math, inspiring messages, and innovative teaching*. Jossey-Bass.

Borich, G. A. (2014). *Effective teaching methods: Research-based practice* (8th ed.).

Brookhart, S. M. (2013). *How to assess higher-order thinking skills in your classroom*. ASCD.

Brookhart, S. M. (2017). *How to give effective feedback to your students*. ASCD.

Brown, G. T. L. (2018). *Assessment of student achievement*. Routledge.

Bulunuz, M., Kıryak, Z., Tomaç, B., Karagöz, F., & Recepoğlu, B. (2017). Evaluation of formative assessment-based teaching practices: An action research. *Turkish Journal of Teacher Education 6(2),* 115-123. http://tujted.com/makale/630

Campbell, D. T., Stanley, J. C., & Gage, N. L. (1963). *Experimental and quasi-experimental designs for research.* Houghton, Mifflin and Company.

Cavanagh, A. J., Aragón, O. R., Chen, X., Couch, B., Durham, M., Bobrownicki, A., Hanauer, D. I., & Graham, M. J. (2016). Student buy-in to active learning in a college science course. *CBE Life Sciences Education, 15*(4), ar76. https://doi.org/10.1187/cbe.16-07-0212

Chappuis, J. (2009). *Seven strategies of assessment for learning.* Pearson.

Clarke, S. (2005). *Formative assessment in the secondary classroom*. Hodder Education.

Collins, N. M. (2012). *The impact of assessment for learning: Benefits and barriers to student achievement* [Unpublished doctoral dissertation]. Cardinal Stritch University.

Creswell, J. W. (2014). *Research design: Qualitative, quantitative, and mixed methods approaches*. Sage.

Earl, L. M. (2003). *Assessment as learning: Using classroom assessment to maximize student learning*. Corwin.

Fennell, F. (2011). *Achieving fluency: Special education and mathematics*. National Council of Teachers of Mathematics.

Gedikli, H. (2018). *The effects of formative assessment practices in science education on students' metacognitive knowledge and skills* (Publication No. 503450) [Master's thesis, Sivas Cumhuriyet University]. Council of Higher Education Thesis Center.

Hattie, J. (2009). *Visible learning: A synthesis of over 800 meta-analyses related to achievement.* Routledge.

Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research, 77*(1), 81-112. https://doi.org/10.3102/003465430298487

Heritage, M. (2007). Formative assessment: What do teachers need to know and do?. *Phi Delta Kappan*, *89*(2), 140-145. https://kappanonline.org/formative-assessment-heritage/

Heritage, M. (2010). *Formative assessment: Making it happen in the classroom*. Corwin.

Heritage, M., Kim, J., Vendlinski, T. P., & Herman, J. L. (2009). From evidence to action: A seamless process in formative assessment?. *Educational Measurement, 28*(3), 24–31. https://doi.org/10.1111/j.1745-3992.2009.00151.x

Hodgson, C., & Pyle, K. (2010). *A literature review of assessment for learning in science*. National Foundation for Educational Research.

Hounsell, D., McCune, V., Hounsell, J., & Litjens, J. (2008). The quality of guidance and feedback to students. *Higher Education Research and Development, 27*(1), 55-67. https://doi.org/10.1080/07294360701658765

Huisman, B., Saab, N. van den Broek, P., & van Driel, J. (2019). The impact of formative peer feedback on higher education students' academic writing: A meta-analysis. *Assessment & Evaluation in Higher Education, 44*(6), 863-880, https://doi.org/10.1080/02602938.2018.1545896

Kıncal, R. Y., & Ozan, C. (2018). Effects of formative assessment on prospective teachers' achievement, attitude and self-regulation skills. *International Journal of Progressive Education, 14*(2), 77-92. https://doi.org/10.29329/ijpe.2018.139.6

King, M. D. (2003). *The effects of formative assessment on student self-regulation, motivational beliefs, and achievement in elementary science* [Unpublished doctoral dissertation]. George Mason University.

Kluger, A. N., & DeNisi, A. (1996). The effects of feedback interventions on performance: A historical review, a meta-analysis, and a preliminary feedback intervention theory. *Psychological Bulletin, 119*(2), 254-284. https://doi.org/10.1037/0033-2909.119.2.254

Koçak, B. (2021). *Determining the attitudes, perceptions and practices of geography teachers towards formative assessment* (Publication No. 676619) [Master's thesis, Necmettin Erbakan University]. Council of Higher Education Thesis Center.

Kramarski, B., & Michalsky, T. (2009). Investigating preservice teachers' professional growth in self-regulated learning environments. *Journal of Educational Psychology, 101*(1), 161–175. https://doi.org/10.1037/a0013101

Kültür, Y. Z. (2021). *The effect of formative assessment on secondary school students' mathematics course achievement and attitudes* (Publication No. 668735) [Master's thesis, Çukurova University]. Council of Higher Education Thesis Center.

Lee, C. (2006). *Language for learning mathematics.* Open University.

McMillan, J. H., & Hearn, J. (2008). Student self-assessment: The key to stronger student motivation and higher achievement. *Educational Horizons, 87*(1), 40-49. https://www.jstor.org/stable/42923742

MoNE, (2017). *Ortaokul matematik dersi (1.,2., 3., 4., 5, 6, 7 ve 8. sınıflar) öğretim programı*.

Nicol, D., & Macfarlane-Dick, D. (2006). Formative assessment and self-regulated learning: A model and seven principles of good feedback practice. *Studies in Higher Education, 31*, 199-218. https://doi.org/10.1080/03075070600572090

Ozan, C., & Kıncal. R. Y. (2018). The effects of formative assessment on academic achievement, attitudes toward the lesson, and self-regulation skills. *Educational Sciences-Theory & Practice, 18*, 85-118. https://doi.org/10.12738/estp.2018.1.0216

Önal, N. (2013). A study on the development of a middle school students' attitudes towards mathematics scale. *Elementary Education Online, 12*(4), 938-948. https://ilkogretim-online.org/index.php/pub/article/view/6042

Pajares, F., & Graham, L. (1999). Self-efficacy, motivation constructs, and mathematics performance of entering middle school students. *Contemporary Educational Psychology, 24*(2), 124-139. https://doi.org/10.1006/ceps.1998.0991

Panadero, E. (2018). Is it safe? Social, interpersonal, and human effects of peer assessment: A review and future directions. In G. T. L. Brown, & L. R. Harris (Eds.), *Handbook of human and social conditions in assessment* (pp. 247-266). Routledge.

Popham, W. J. (2008). *Transformative assessment*. ASCD.

Popham, W. J. (2011). *Classroom assessment: What teachers need to know* (6th ed.). Allyn & Bacon.

Razali, N., & Wah, Y. (2011). Power comparisons of Shapiro-Wilk, Kolmogorov-Smirnov, Lilliefors and Anderson-Darling tests. *Journal of Statistical Modeling and Analytics, 2*(1), 21-33. http://www.de.ufpb.br/~ulisses/disciplinas/normality_tests_comparison.pdf

Rodriguez, M. C., & Albano, A. D. (2017). *The college instructor's guide to writing test items: Measuring student learning*. Routledge.

Ruiz-Primo, M. A., Briggs, D., Iverson, H., Talbot, R., & Shepard, L. A. (2011). Impact of undergraduate science course innovations on learning. *Science*, *331*, 1269-1270. https://doi.org/10.1126/science.1198976

Sadler, D. R. (1989). Formative assessment and the design of instructional systems. *Instructional Science, 18*(2), 119-144. https:/doi.org/10.1007/BF00117714

Sadler, D. R. (1998). Formative assessment: revisiting the territory. *Assessment in Education: Principles, Policy & Practice*, *5*(1), 77-84. https://doi.org/10.1080/0969595980050104

Sadler, P. M., & Good, E. (2006). The impact of self- and peer-grading on student learning. *Educational Assessment*, *11*(1), 1-31. https://doi.org/10.1207/s15326977ea1101_1

Schiefele, U., Schaffner, E., Möller, J., & Wigfield, A. (2012). Dimensions of reading motivation and their relation to reading behavior and competence. *Reading Research Quarterly*, *47*(4), 427-463. https://doi.org/10.1002/rrq.030

Schoenfeld, A. H. (2013). *Mathematical problem solving*. Elsevier.

Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference.* Houghton, Mifflin and Company.

Shavelson, R. J., Young, D. B., Ayala, C. C., Brandon, P. R., Furtak, E. M., Ruiz-Primo, M. A., Tomita, M. K., & Yin, Y. (2008). On the impact of curriculum-embedded formative assessment on learning: A collaboration between curriculum and assessment developers. *Applied Measurement in Education*, *21*(4), 295-314. https://doi.org/10.1080/08957340802347647

Shute, V. J. (2008). Focus on formative feedback. *Review of Educational Research, 78*(1), 153-189. https://doi.org/10.3102/00346543073137

Sönmez, M. (2020). *The effect of formative assessment on reading comprehension* (Publication No. 649617) [Master's thesis, Düzce University]. Council of Higher Education Thesis Center.

Swan, M. (2006). *Collaborative learning in mathematics: A challenge to our beliefs and practices*. National Institute of Adult Continuing Education.

Tabachnick, B. G., & Fidell, L. S. (2019). *Using multivariate statistics* (7th ed.). Pearson.

Topping, K. J. (2009). Peer assessment. *Theory into Practice, 48*(1), 20-27. https://doi.org/10.1080/00405840802577569

Tuominen, K. R. (2008). *Formative assessment and collaborative teaming with support involving middle school mathematics teachers* [Unpublished doctoral dissertation]. California University.

Unaş, F. (2021). *The effects of formative assessment practices in science education on students' self-regulated learning skills* (Publication No. 693052) [Master's thesis, Sivas Cumhuriyet University]. Council of Higher Education Thesis Center.

Walsh, J. A. (2022). *Questioning for formative feedback: Meaningful dialogue to improve learning*. ASCD.

Walsh, J. A., & Sattes, B. D. (2016). *Questioning for classroom discussion: Purposeful speaking, engaged listening, deep thinking*. ASCD.

Wigfield, A., & Eccles, J. S. (2000). Expectancy–value theory of achievement motivation. *Contemporary Educational Psychology, 25*(1), 68-81. https://doi.org/10.1006/ceps.1999.1015

Wiliam, D. (2011). What is assessment for learning?. *Studies in Educational Evaluation, 37*(1), 3-14. https://doi.org/10.1016/j.stueduc.2011.03.001

Wiliam, D. (2018). Feedback: At the hard of -but definitely not all of- formative assessment. In A. A. Lipnevich, & J. K. Smith (Eds.), *The Cambridge handbook of instructional feedback* (pp. 3-28). Cambridge University.

Yin, Y., Shavelson, R. J., Ayala, C. C., Ruiz-Primo, M. A., Brandon, P. R., Furtak, E. M., Tomita, M. K., & Young, D. B. (2008). On the impact of formative assessment on student motivation, achievement, and conceptual change. *Applied Measurement in Education, 21*(4), 335-359. https://doi.org/10.1080/08957340802347845

Necatibey Faculty of Education Electronic Journal of Science and Mathematics Education
Vol. 18, Issue 2, December 2024, pp. 375-403. ISSN: 1307-6086

Research Article

# Assessing the Pre-Service Science and Mathematics Teachers' Systems Thinking Skills through Case Scenarios

**Gaye Defne CEYHAN[1], Ulku Seher BUDAK[2], Busra KARGA[3]**

[1] Bogazici University, Department of Mathematics and Science Education, Türkiye, gaye.ceyhan@bogazici.edu.tr, http://orcid.org/0000-0003-1312-3547

[2] Bogazici University, Department of Mathematics and Science Education, Türkiye, ulkuseherbudak@gmail.com, https://orcid.org/0000-0002-4047-9920

[3] Eindhoven University of Technology, Department of Applied Science and Physics Education, The Netherlands, busrakarga@gmail.com, https://orcid.org/0000-0001-5356-8482

Received : 07.07.2024 Accepted : 01.11.2024

Doi: https://doi.org/10.17522/balikesirnef.1511926

---

*Abstract* – Addressing complex global problems requires more comprehensive and holistic approaches that highlight the necessity of systems thinking skills; however, existing studies indicate a significant gap in understanding the systems thinking skills of pre-service teachers, emphasizing the need for further research in this area. This exploratory case study research explored the systems thinking skills of pre-service science and mathematics teachers through scenario-based assessments. Three case scenario examples focused on a specific aspect of systems thinking: stock-flow, causal-loop, and dynamic thinking. The participants of this study were 14 pre-service teachers taking a systems thinking course at the teacher education program of a public research university. The data were coded using the Systems Thinking Rubric and the Dynamic Thinking Skills Rubric. The results revealed that participants made notable improvements in dynamic thinking. However, fewer participants exhibited growth in stock-flow thinking comparing the participants' disciplines, the results showed that pre-service science teachers demonstrated greater advancements in systems thinking skills than their mathematics counterparts. This exploratory research offers insights into assessing systems thinking skills in pre-service teachers. Integrating a systems thinking approach into teacher training programs could enhance teachers' preparedness to comprehend complex issues. Further studies employing systems thinking practices in teacher training programs could elucidate the optimal development of systems thinking among aspiring teachers. Therefore, this research demonstrates the potential of systems thinking to enrich pre-service teacher education.

*Keywords*: Pre-service teachers, systems thinking assessment, systems thinking skills, teacher learning.

------------

Corresponding author: Gaye Defne CEYHAN, gaye.ceyhan@bogazici.edu.tr

## Introduction

Understanding complex issues may require a systems thinking approach (Forrester, 1994) because systems thinking may help individuals see how different components within a system interact and potentially influence each other (Ben-Zvi Assaraf & Orion, 2010; Evagorou et al., 2009). Developing systems thinking skills (STS) is crucial for identifying and interpreting the systems in the interconnected world we live in (Bielik et al., 2023; Hmelo-Silver & Pfeffer, 2004; Sweeney & Sterman, 2007). Integrating systems thinking into education may enhance individuals' comprehension of systems and hopefully manage real-world challenges effectively (Fisher & Systems Thinking Association, 2023; Senge, 1990; Eidin et al., 2023). Therefore, pre-service teachers (PSTs) need to develop strong STS to navigate the complexities of learning and teaching the topics in their field (Ateskan & Lane, 2018; Karaarslan Semiz & Teksöz, 2020; Lee et al., 2019). Assessing these skills can help teacher educators determine the preparedness of PSTs and highlight areas that may need further improvement (Lavi & Dori, 2019; Lee et al., 2019).

Systems thinking focuses on the interrelationships of parts of a system, while system dynamics utilizes conceptual or mathematical models to simulate these interactions over time (Senge, 1990; Sweeney & Sterman, 2000). Systems thinking and system dynamics provide frameworks that may make it easier to understand how systems work (Arnold & Wade, 2015). Within these frameworks, stock flow thinking, dynamic thinking, and closed-loop thinking are considered important skills in systems thinking (Dorani et al., 2015). Stock-flow thinking involves recognizing and differentiating system quantities (stocks) and the rates at which they change (flows) (Aşık & Doğanca Küçük, 2021; Sweeney & Sterman, 2000). Causal loop thinking is about determining feedback loops, where changes in one part strengthen or balance the changes within the system (Doganca Kucuk & Saysel, 2018). Dynamic thinking includes perceiving the system from a long-term perspective, which helps the learner anticipate the system's development and adaptation over time, usually in response to internal and external shifts (Dorani et al., 2015; Richmond, 1993).

Researchers emphasize the significance of developing skills like comprehending stock-flow, causal-loops, and dynamic thinking for the management of systems within environments (Hopper & Stave, 2008; Plate, 2010). These skills enable individuals to navigate systems and make decisions to generate expected outcomes. However, assessing STS has been challenging (Lee et al., 2019). The complexity and dynamic nature of systems pose challenges in assessing these skills (Fisher & Systems Thinking Association, 2023).

Conventional assessment methods, such as multiple-choice tests, are often considered inadequate in effectively evaluating an individual's STS (Dorani et al., 2015). Moreover, there is no consensus among experts and educators regarding the definition and structure of systems thinking, which complicates the development of assessments of STS (Hopper & Stave, 2008; Lee et al., 2019; Stave & Hopper, 2007).

An individual's STS may be revealed by using scenarios to analyze how systems solve problems (Dorani et al., 2015). A person's comprehension and use of STS may be reflected in case-based scenarios. We can develop more effective methods of evaluating STS by integrating approaches and resolving evaluation flaws (Lee et al., 2019). This study uses three case scenarios from Dorani et al. (2015) to examine the STS of PSTs. The evaluation of pre-service teachers' systems thinking skills is of paramount importance as it aids in confirming their readiness to educate in an intricate and interlinked global context (Dorani et al., 2015). By equipping them with systems thinking skills, teachers can lay the groundwork for their students' success across various sectors, from the sciences and engineering to business and politics (Peretz et al., 2023).

Evaluating pre-service teachers' systems thinking skills can assist in identifying areas requiring further training or support (Karaarslan Semiz & Teksöz, 2020) and promote their ability to impart systemic perspectives to their future students (Kriswandani et al., 2022). Our research contributes to existing literature by offering examples for teacher education programs to promote STS development and investigating how a course focused on systems thinking impacts PSTs' stock flow, causal loop, and dynamic thinking skills. This investigation deepens our knowledge of how targeted interventions can nurture these skills in educators potentially enhancing their ability to navigate complex educational systems and make well-informed decisions in their professional roles.

The primary research question guiding this study is: How does a systems thinking-oriented teacher education course change PSTs' levels of stock-flow, causal loop, and dynamic thinking skills? The study addresses the following sub-questions:

(1) How do pre-service teachers' levels of stock-flow thinking change from pre-survey to post-survey?

(2) How do pre-service teachers' levels of causal loop thinking change from pre-survey to post-survey?

(3) How do pre-service teachers' levels of dynamic thinking change from pre-survey to post-survey?

This study aims to add to the expanding body of work on systems thinking in teacher training and provide ideas for creating teacher preparation programs that nurture these skills. By investigating these inquiries, our goal is to illuminate the process by which future teachers acquire STS and how teacher training programs can be improved to offer assistance in this area. The results could shape strategies for incorporating systems thinking into fields, within teacher education ultimately improving the thoroughness and effectiveness of preparing future educators to address intricate classroom issues.

**Systems Thinking in Education**

Understanding how different parts of a system interact and influence each other is essential in today's world (Forrester, 2007; Peretz et al., 2023). This way of thinking is especially important in education, because it may help future teachers grasp the dynamics within classrooms and improve the effectiveness of their teaching on complex issues (Karaarslan Semiz & Teksöz, 2020; Kriswandani et al., 2022; Uskola & Puig, 2022). Systems thinking in education also offers a way to perceive issues from various perspectives, which may empower teachers to create valuable learning experiences for their students (Boubonari et al., 2023; Mehren et al., 2018).

Educators use systems thinking to identify and address topics in various fields of education (Fisher & Systems Thinking Association, 2023; Meadows, 2008). This approach covers fields, such as science, technology, engineering, and mathematics (STEM), social sciences, and the arts, by integrating topics and relating them to real-world situations (Ben-Zvi Assarraf et al., 2013; York et al., 2019). Moreover, it boosts creativity and innovation by taking into account viewpoints (York et al., 2019). However, despite its advantages, the implementation of systems thinking, in education, is constrained by a lack of comprehension and resources (Fisher & Systems Thinking Association, 2023).

The incorporation of systems thinking into K-12 education faces obstacles because teachers may not be adequately prepared (Bartus & Fisher, 2016; Taylor et al., 2020). While educators often excel in skills, they may lack the know-how needed for systems thinking (Bartus & Fisher, 2016). Experts stress the importance of having a systems thinking framework in curriculum design as cross-disciplinary approaches can enhance education on a scale (Peretz et al., 2023). Many teachers lack training in this area, which hinders their effectiveness, in teaching and understanding systems.

Courses that focus on Pedagogical Content Knowledge (PCK) have proven to be more effective, in teaching systems thinking, compared to those that prioritize methods (Rosenkränzer et al., 2017). Professional development is crucial in aiding teachers in instructing systems and using models in science education (Yoon et al., 2017). PCK involves teaching systems through strategies rather than just definitions (Peretz et al., 2023; Rosenkränzer et al., 2017). A course on systems thinking does not have to cover all skills at once; assessments should align with objectives (Plate & Monroe, 2014).

This research explored the changes in PSTs' skills in stock flow, causal loop, and dynamic thinking from pre- to post-survey. This research aimed to contribute to the study of teaching and evaluating systems thinking in teacher education. Assessing the STS of PSTs is crucial for investigating their competencies to teach complex topics effectively (Dorani et al., 2015). Such evaluations assist in identifying areas that need improvement and supporting PSTs' skills to enhance their perspectives (Karaarslan Semiz & Teksöz, 2020; Kriswandani et al., 2022).

**Assessment of Systems Thinking Skills**

Various frameworks exist for evaluating STS in education (Arnold & Wade, 2015; Ben-Zvi Assarraf et al., 2005; Stave & Hopper, 2007). Each offers distinct approaches to defining and assessing systems thinking. Ben-Zvi Assarraf and colleagues (2005) categorized skills based on proficiency levels in the Systems Thinking Hierarchy (STH) model. The STH model emphasizes three core skills: (1) identifying the system's components and processes, (2) recognizing relationships among components, and (3) understanding the dynamic relationships within the system. Using similar levels in the STH model, Stave and Hopper (2007) developed a taxonomy including seven STS and identified commonly cited characteristics of systems thinking in the literature as distinguishing between flows and variables, understanding dynamic behavior, and recognizing feedback. Arnold and Wade (2015) introduced a framework that focuses on problem definition and system analysis.

Our research uses the STH model as a framework because the STH framework offers a comprehensive and structured approach to evaluating systems thinking by emphasizing the three core skills. These fundamental abilities are in line with the goal of our study, which is to evaluate how well teacher candidates understand the interdependencies and dynamic behaviors that are essential for climate teaching in addition to identifying system components. Moreover, the STH framework's focus on building from basic identification to understanding complex system interactions provides a developmental approach to assessing how individuals engage with systems thinking over time. Finally, the three core skills of the STH model are critical for comprehending climate systems, which are inherently complex and interconnected.

Assessment of individuals' STS includes diverse techniques including concept maps, questionnaires, interviews, case-based real-world scenarios, and observations (Ateskan & Lane, 2018; Batzri et al., 2015; Evagorou et al., 2009). Scenarios are used to evaluate learners' causal reasoning skills (Dorani et al., 2015). Detailed rubrics aligned with these frameworks help standardize the evaluation processes (Lee et al., 2019). Integrating assessment tools, with established frameworks, could improve the evaluation of STS (Budak & Ceyhan, 2024; Karaarslan Semiz & Teksöz, 2020). While challenges exist, utilizing qualitative and performance-based approaches has provided insights into the STS of individuals.

Assessing stock flow, causal relationships, and dynamic thinking may provide insights into individuals' overall understanding of systems. Evaluations often involve scenario-based tasks where learners explain the interaction between variables, create diagrams, predict system states, and consider delays. Developing these interconnected skills may equip individuals to effectively interpret and analyze real-world system complexities (Evagorou et al., 2009). Research on systems thinking has been conducted with a variety of participants, including elementary, secondary, and higher education; of fields, including science, mathematics, and engineering; and topics including sustainability and healthcare (Peretz et al., 2023). Peretz and colleagues (2023) stated that cross-disciplinary curricula that encourage systems thinking may be beneficial for science and engineering courses as well as for education in general.

The systems thinking approach has been studied in science education on biology (Ben-Zvi Assaraf et al., 2013; Riess & Mischo, 2010; Sommer & Lücken, 2010; Tripto et al., 2018); on chemistry (Delaney et al., 2021; Eaton et al., 2019); on earth science (Ben-Zvi Assaraf & Orion, 2005; 2010; Evagorou et al., 2009; Kali et al., 2003; Lee et al., 2019; Mehren et al., 2018); on Education for Sustainable Development (ESD) (Ateskan & Lane, 2018; Doganca Kucuk & Saysel, 2018; Karaarslan Semiz & Teksöz, 2020; Meilinda et al., 2018); on physics (Nuhoğlu, 2010). On the other hand, studies conducted with students, teachers, or teacher candidates in the field of mathematics education are limited (Salado, et al., 2019). Research in the literature reveals that different participant groups have exhibited a variety of skills across various aspects of systems thinking. These studies indicated that students' systems thinking skills are limited in understanding systems, highlighting the need for a systems thinking approach in teaching these subjects (Karga & Ceyhan, 2024).

## Method

This study used qualitative research methodology, which allows for an in-depth understanding of the phenomenon under investigation (Gay et al., 2012). Specifically, an exploratory case study design was used in this study, which sought to answer the how and why of the phenomenon under investigation (Yin, 2009). Three case-based scenarios were implemented, each focusing on stock flow, causal loop, or dynamic thinking to evaluate the STS of PSTs. Participants were asked to fill out surveys before and after the course implementation where they were asked to answer open-ended questions based on case studies validated by Dorani et al. (2015). The researchers delve into the intricacies of how participants perceive science and math education by analyzing their written responses using predefined evaluation criteria and coding systems.

### Context of the Study and Participants

This research was carried out at a public research university in Türkiye, with a reputation for offering undergraduate and graduate programs in various fields such as engineering, applied sciences, and social sciences. Purposive sampling is when the researcher deliberately selects a sample by identifying criteria (Gay et al., 2012). In this study, purposive sampling was used because the effect of the systems thinking course on the participants' skills was to be measured. The study focused on science and math educators in the Mathematics and Science Education Department who were taking a course on systems thinking in science and math education. The group of participants included 14 PSTs. Nine were from the science education program and five were from the math education program. These individuals were preparing to teach middle school students (grades five to eight). Among them, there were 13 female participants and one male participant. Twelve of them were nearing completion of their studies, while two were starting their year of a master’s program focusing on science and math education. Their ages ranged from 22 to 25 years, with a mean age of 23.7 years (standard deviation = 1.26).

The elective course undertaken by study participants served as an introduction to systems thinking, tailored specifically for Mathematics and Science Education contexts.

Throughout the course, participants explored the definition and practical application of systems thinking processes. They received a comprehensive overview of fundamental concepts and systems dynamics tools. The curriculum provided a broad introduction to the systems thinking approach, examining in depth the characteristics of complex systems as outlined in Table 1. This course aimed to equip PSTs with the conceptual framework and analytical tools necessary to apply systems thinking principles in educational settings. By focusing on applications relevant to mathematics and science instruction, the course sought to enhance participants' skills to navigate and address multifaceted classroom challenges using a systems-oriented perspective.

**Table 1** Flow of the Systems Thinking in Science and Mathematics Education Course

| Weeks | Content |
|---|---|
| Week 1-2 | Introduction to systems thinking approach and its relevance in science & math education (pre-survey) |
| Week 3 | Behavior over time graphs and its applications in science & math education |
| Week 4-5 | Stock & flow diagrams and its applications in science & math education |
| Week 5-6 | Causal loop diagrams and its applications in science & math education |
| Week 7-8 | Dynamic thinking and its applications in Science & Math Education |
| Week 9 | Exploring systems thinking tools in lesson plans |
| Week 10-11 | Modeling systems in education (STELLA) |
| Week 12-13 | Developing lesson plans with systems thinking tools |
| Week 14 | Wrap up the course (post-survey) |

As shown in Table 1, the course also included integrating systems thinking principles into instructional design. The intention was to harmonize traditional pedagogical techniques with the holistic outlook of systems thinking. Additionally, the course advocated for implementing inquiry-based instructional strategies that align with the systems thinking approach. This approach highlights the significance of active participation and asking questions, which helps promote essential skills needed to understand and navigate the difficulties of complex system dynamics.

**Procedure**

This research used surveys before and, after a semester to assess changes in participants' STS (refer to Table 1). Three scenario-based cases, which were developed by Dorani et al. (2015), were used to assess participants' stock flow, causal loop, and dynamic thinking skills. The university's ethics committee approved the study and participants gave their consent to participate in the study. The research included three scenarios representing varying levels of system complexity. The first case illustrated a system with a minimal number of variables. The second case presented a system with multiple variables and feedback loops. The third case, the one depicted a dynamic system with numerous feedback loops and nonlinear relationships between variables. By using these scenarios, the research aimed to assess how well PSTs could apply systems thinking across levels of complexity.

## Data Collection

Different data collection tools have been used in the literature to evaluate participants' systems thinking skills, including multiple-choice and skill-based tests, questionnaires, concept maps, and interview questions (Ben-Zvi Assaraf & Orion, 2005; Lee et al., 2019; Mambrey et al., 2020). However, Dorani's (2015) context-independent scenarios were chosen as a data collection tool because context-independent scenarios ensure that assessment outcomes are not affected by teachers' specific content knowledge, thus offering a more accurate evaluation of their overall systems thinking skills (Karga & Ceyhan, 2024). It was also noted that scenario-based questions are arguably better at assessing participants' potential behavior in realistic situations (Daniel & Mazzurco, 2020).

### *The Case on Stock-Flow Thinking*

A stock-flow thinker can distinguish stocks and flows. This individual understands that changes in stock variables can only be achieved indirectly through adjustments in flow variables with a delay. To illustrate this concept, the study referenced a scenario where a city dealt with a rat infestation (Dorani et al., 2015). Despite implementing a temporary resolution by deploying rat poison, the problem reemerges due to disregarding the rat birth rate. A possible reply might mention the factor of the disrupted balance, where the high rate of rat reproduction has caused an increase in the rat community.

### *The Case of Causal Loop Thinking*

A causal loop thinker recognizes that every action and decision can trigger unexpected outcomes, which subsequently shape the context for future decision-making. The scenario-based example of a farming village that opted to use a pesticide to combat a green bug infestation, not realizing that these green bugs also prey on detrimental red bugs, was used in this study (Dorani et al., 2015). Despite their efforts, the villagers continued to encounter crop damage. This unforeseen outcome was the surge in the population of red bugs due to the elimination of their natural predators. Participants were anticipated to discern that the decision to exterminate the green bugs has indirectly reshaped the future state of the problem, resulting in an escalation in the red bug population.

### *The Case of Dynamic Thinking*

By definition, a dynamic thinker can discern incremental shifts and accurately recognize trends and behavior patterns over in a while. A question that can effectively evaluate this skill might present a decision-making scenario with two choices. In this study, the research focused on a situation where a person was comparing two real estate investments; a property, in a desirable city area and a bigger house, in a growing but affordable neighborhood (Dorani et al., 2015). An appropriate response would involve explaining the effects of different stakeholder perspectives, gradual changes, potential feedback loops, and delays on systemic behavior.

## Data Analysis

Two rubrics were used to analyze participants' responses to the three cases. The Systems Thinking Rubric, developed by Lee et al. (2019), was used for the first two cases: stock-flow and causal loop thinking. Lee et al. (2019) used an inductive approach to develop the rubric to define different levels of STS and explore their application in proposed lessons on the water cycle.

The rubric manifested four distinct proficiency levels, with each participant's response accordingly coded to a specific level. The categorization of these levels was founded on an assessment of participants' responses, utilizing components of the Systems Thinking Hierarchical (STH) Model (Ben-Zvi Assaraf & Orion, 2005), interrelationships among subsystem processes and components, as well as the Next Generation Science Standards (NGSS) crosscutting concepts (NGSS, 2013). Lee et al. (2019) named the four levels novice, recognition, beginning, and intermediate. Regarding the advancement of the levels and to be aligned with the other rubric used in this study, the names of the four levels were revised as novice, developing, intermediate, and advanced (Table 2).

**Table 2** The Systems Thinking Rubric (Adapted from Lee et al., 2019)

| Level | Description |
|---|---|
| Novice (Level 0) | Lack of response or an explicit implication of unfamiliarity with the given system |
| Developing (Level 1) | Identifying a single part, process, or pattern within the system |
| | Lack of elaboration on the relationship between the parts and processes |
| Intermediate (Level 2) | Identifying at least two parts or processes |
| | Limited to one-directional cause and effect (e.g., A causes B) or recognizing a relationship solely between two components |
| Advanced (Level 3) | Identifying three or more parts or processes, with an understanding that involves at least two or more interacting parts |
| | Multiple interactions are recognized, demonstrating an increased complexity in understanding the system |

The third case assessed participants' dynamic thinking skill levels using the revised rubric developed by the authors (Karga & Ceyhan, 2024). The rubric was developed based on the revised System Thinking Rubric (Lee et al., 2019) to ensure consistency when analyzing data from the different scenarios. In developing the Dynamic Thinking Skills Rubric, an expert view was obtained regarding content coverage, criteria selection, and descriptor clarity (Karga & Ceyhan, 2024). The level descriptions of the Dynamic Thinking Skills Rubric are given in Table 3.

**Table 3** The Dynamic Thinking Skills Rubric

| Level | Description |
|---|---|
| Novice (Level 0) | No comprehension or application of dynamic thinking concepts |
| | Decisions are made based solely on personal preferences or immediate costs |
| Developing (Level 1) | Demonstrating an understanding of behavioral patterns within a system or over time, with consideration of short-term or temporary factors |
| Intermediate (Level 2) | Expanding the mental models to include past and future trends, with an awareness of potential growth or stability |
| Advanced (Level 3) | Explaining how different stakeholder views, gradual changes, feedback loops, and delays affect overall system behavior |

Two of the researchers independently coded each participant's pre- and post-survey responses to the three cases using the Systems Thinking Rubric (Lee et al., 2019) for stock-flow and causal-loop thinking, and the Dynamic Thinking Skills Rubric (Karga & Ceyhan, 2024) for dynamic thinking. The first researcher has a Ph.D. in science and has conducted various research on systems thinking. The second researcher has a master's degree in science education and specializes in systems thinking in science education. Interrater agreement was calculated using the formula: # agreements / (# agreements + # disagreements) x 100 (Cooper et al., 2019). The researchers met and compared their scores for each participant, with initial agreement rates of .75 for stock-flow thinking, .71 for causal loop thinking, and .86 for dynamic thinking. Then, the researchers explained to each other the logic of the codes that they had done differently. They discussed their conflicts until they reached a complete agreement across their codes.

## Results

To answer the research question in more detail, under three sub-research questions, this study examined PSTs' levels of stock-flow thinking, causal loop thinking, and changes in dynamic thinking from pre-survey to post-survey. The results of the sub-questions were presented in three categories under this heading.

### How do pre-service teachers' levels of stock-flow thinking change from pre-survey to post-survey?

The pre-survey to post-survey levels of PSTs' stock-flow systems thinking for Case 1 are shown in Figure 1. When the pre-survey responses of fourteen participants were analyzed to determine the stock-flow thinking levels, two participants were at Level 1, eleven were at Level 2, and one was at Level 3. When the participants' responses to the post-survey were analyzed, it was found that five participants were at Level 1, four were at Level 2, and five were at Level 3. No Level 0 respondents were recorded in either the pre- or post-survey.

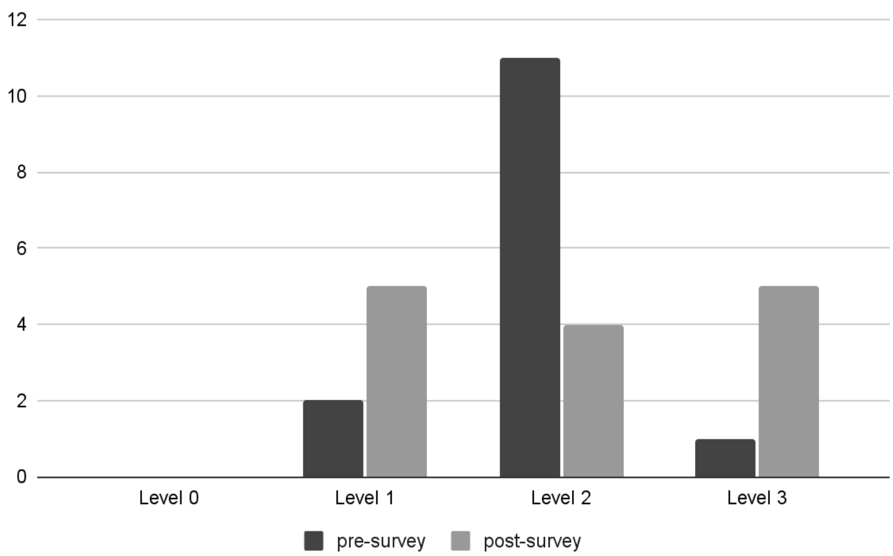



**Figure 1** Participants' Pre- and Post-survey Stock-flow Thinking Levels

Table 4 shows the sample quotes given by participants for Case 1 on the pre-and post-surveys.

**Table 4** Participants' Stock-flow Thinking Levels and Sample Quotes for Case 1

| Levels | Example quotes for pre-survey | Example quotes for post-survey |
|---|---|---|
| Level 0 | - | - |
| Level 1 | "The rest of the rats that survived can breed new ones. And as I know, they give birth to more than one at a time." (M3) | "I think the rats can become resistant to the poison." (S1) |
| Level 2 | "A few mice that did not die may have developed resistance and multiplied due to the poison placed in the environment." (S6) | "Since the mice did not all die at the same time, the remaining mice may have developed a defense against the poison over time. Mice unaffected by the poison may have reproduced." (M4) |
| Level 3 | "Pollution is seen more in tourist places due to population. Since rats generally live in dirty environments, it causes the mouse population to reappear. Since rat populations are not common in a clean environment, the main solution should be to give importance to environmental cleanliness." (M5) | "The increase in the number of rat species unaffected by the poison over time led to an increase in the number of rats in the city. The fact that the remaining species now ate the food sources of the poisoned mice and that they could easily find food, may have contributed to this. Eliminating just one of the environmental conditions does not solve the whole problem." (S5) |

S stands for science PSTs, and M stands for mathematics PSTs. "-" stands for no answer.

### How do pre-service teachers' levels of causal loop thinking change from pre-survey to post-survey?

Figure 2 shows the participants' pre- and post-survey causal loop thinking levels for Case 2. When the pre-survey responses of a total of fourteen participants are analyzed to determine the participants' causal thinking levels, it can be seen that one of the participants was at Level 0, four were at Level 1, six were at Level 2, and three were at Level 3. When the participants' responses to the post-survey were analyzed, it was found that three participants were at Level 2 and eleven were at Level 3. No Level 0 and Level 1 respondents were recorded on the post-survey.

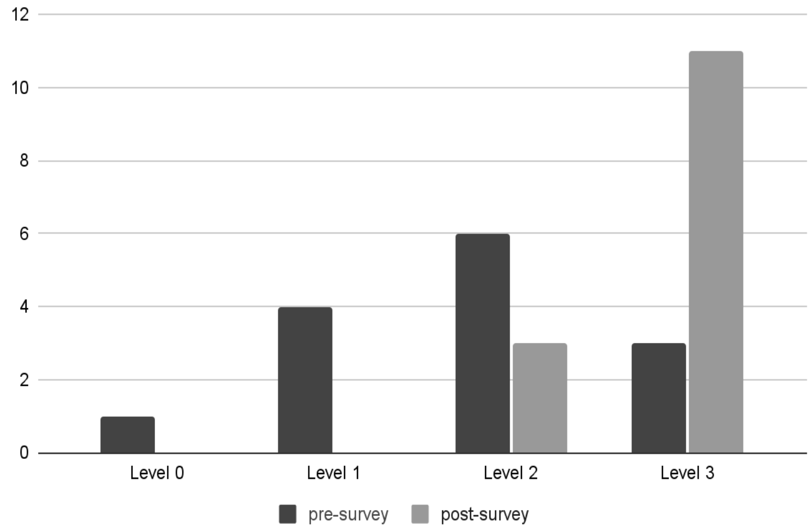



**Figure 2** Participants' Pre- and Post-survey Causal Loop Thinking Levels

Table 5 shows the sample quotes given for Case 2 on the pre-and post-surveys.

**Table 5** Participants' Pre-and-post Surveys Causal Loop Thinking Levels and Sample Quotes for Case 2

| Levels | Example quotes for pre-survey | Example quotes for post-survey |
|---|---|---|
| Level 0 | "Because they broke the system. Even though they think they have found a solution, it has created new problems." (S1) | - |
| Level 1 | "The villagers have destroyed the food chains. It is a cycle and it causes problems for every part of the pieces of the system." (S8) | - |
| Level 2 | "Just as in the food web, the extinction of the green insects led to the reappearance of the red insects, and at the same time there were changes in the number of animals, which increased and decreased as the system of the food chain was disrupted." (S5) | "The farmers did not focus on the main problem. They just tried to solve their problem in a way that was more appropriate for them. So they used the pesticide to kill the green bugs. To find the right solution, they have to find out why these things happen. This requires them to look at deeper levels of abstraction within the system that are not immediately obvious." (S4) |
| Level 3 | "The farmer should try to destroy the red bugs so that the population of green bugs will decrease and his crops will remain healthy." (M5) | "Because the farmers had found a short-term solution. They did not consider the future possibilities. For a certain period after using this pesticide, the insect may become resistant, or the plants may be damaged by the pesticide." (S7) |

S stands for science PSTs, and M stands for mathematics PSTs. "-" stands for no answer.

**How do pre-service teachers' levels of dynamic thinking change from pre-survey to post-survey?**

Figure 3 shows the participants' pre-survey and post-survey levels of dynamic thinking for Case 3. When the pre-survey responses of fourteen participants were analyzed to determine the dynamic thinking levels, six participants were at Level 0, seven were at Level 1, one was at Level 2, and no one was at Level 3. When the participants' post-survey responses were analyzed, it was found that two participants were at Level 0, two were at Level 1, nine were at Level 2, and one was at Level 3.

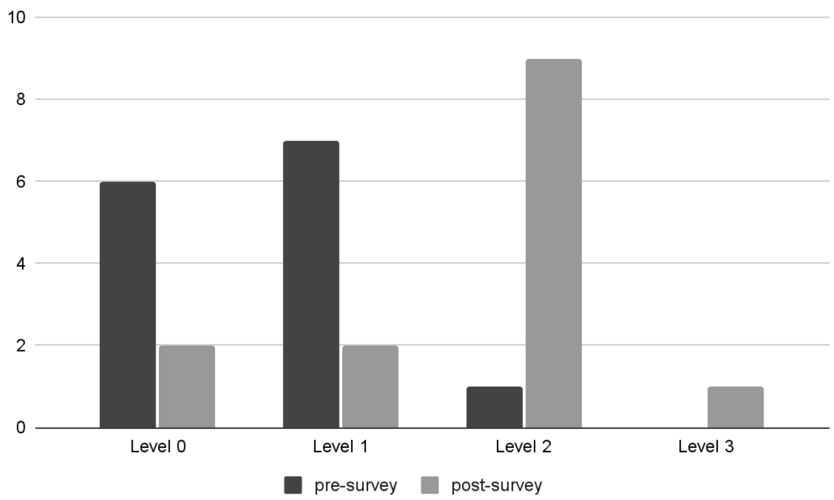



**Figure 3** Participants' Pre-and Post-survey Dynamic Thinking Levels

Table 6 shows the sample quotes given for Case 3 on the pre-and post-surveys.

**Table 6** Participants' Dynamic Thinking Levels and Sample Quotes for Case 3

| Levels | Example quotes for pre-survey | Example quotes for post-survey |
|---|---|---|
| Level 0 | "I would choose a small house near the heart of the city because it is time efficient. Also, there are plenty of opportunities to socialize." (S4) | "If I were Kramer, I would choose a big house because I think a big house is more useful. Also, it is less expensive, and I think the place where the big house is located is quieter than the big house." (S3) |
| Level 1 | "I would choose a large house in a less expensive but up-and-coming neighborhood. The big house is advantageous both in terms of area and budget, and I think it is the right choice because it is already developed." (M1). | "I would prefer the second house because it is mentioned as an emerging neighborhood. So, it not only provides a comfortable living space for not but also has potential to be valuable." (S1) |

**Table 6** (continued)

| Levels | Example quotes for pre-survey | Example quotes for post-survey |
|---|---|---|
| Level 2 | "I choose the second home because I prefer a place that is open to innovation rather than the existing one. Even if it is a cheaper house, I think it can be advantageous if it is bigger than the other one and the surrounding area is also developing." (S5). | "I would choose the big house. because it is both affordable and highly likely to increase in value in the future. and also for investment purposes, so it's better to have a big house." (M1). |
| Level 3 | - | "If I were Kramer, I'd choose the house in the developing area. In the beginning, he will have both a larger and more affordable house, and at the same time, many different factors will change and develop, and he will have the advantage of the location of the first house in the future. Even if it is the case of choosing the first house, it should be taken into account that the conditions of the environment there may also change over time." (S5) |

S stands for science PSTs, M stands for mathematics PSTs. "-" stands for no answer.

This study examined the contribution of a course using inquiry-based teaching strategies compatible with the systems thinking approach to developing STS of PSTs. This study evaluated how participants' stock-flow thinking skills for the first case, causal loop thinking skills for the second case, and dynamic thinking skills for the third case changed from the pre-survey to post-survey across three cases. Table 7 provides a holistic view of the participants' levels by scoring their responses to pre- and post-survey cases. The table shows which systems thinking skill(s), if any, changed for mathematics and science PSTs during the course from the pre- to post-survey.

**Table 7** Pre and Post-survey Results of the Three Cases

| | Case 1: Stock-flow thinking | | Case 2: Causal loop thinking | | Case 3: Dynamic thinking | |
|---|---|---|---|---|---|---|
| | Pre-survey | Post-survey | Pre-survey | Post-survey | Pre-survey | Post-survey |
| **M1** | Level 2 | Level 1 | Level 2 | Level 3 | Level 1 | Level 2 |
| **M2** | Level 1 | Level 1 | Level 3 | Level 3 | Level 0 | Level 2 |
| **M3** | Level 1 | Level 1 | Level 3 | Level 3 | Level 0 | Level 2 |
| **M4** | Level 2 | Level 2 | Level 1 | Level 2 | Level 0 | Level 0 |
| **M5** | Level 3 | Level 3 | Level 3 | Level 3 | Level 1 | Level 2 |
| **S1** | Level 2 | Level 1 | Level 0 | Level 3 | Level 0 | Level 1 |
| **S2** | Level 2 | Level 3 | Level 2 | Level 3 | Level 1 | Level 2 |
| **S3** | Level 2 | Level 1 | Level 1 | Level 2 | Level 0 | Level 0 |
| **S4** | Level 2 | Level 2 | Level 2 | Level 2 | Level 0 | Level 1 |
| **S5** | Level 2 | Level 3 | Level 2 | Level 3 | Level 2 | Level 3 |
| **S6** | Level 2 | Level 3 | Level 2 | Level 3 | Level 1 | Level 2 |
| **S7** | Level 2 | Level 2 | Level 1 | Level 3 | Level 1 | Level 2 |
| **S8** | Level 2 | Level 3 | Level 1 | Level 3 | Level 1 | Level 2 |
| **S9** | Level 2 | Level 2 | Level 2 | Level 3 | Level 1 | Level 2 |

S stands for science PSTs, M stands for mathematics PSTs. "-" stands for no answer.

### *Stock-flow Thinking*

When Table 7 is analyzed in terms of levels of stock-flow thinking, the number of participants at Levels 1 and 3 increased from the pre-survey to the post-survey, while the number of participants at Level 2 decreased. Two participants who were at Level 1 on the pre-survey showed no improvement on the post-survey and remained at Level 1. Eleven participants were at Level 2 on the pre-survey. On the post-survey, three dropped to Level 1, four remained at the same level, and four moved up to Level 3. The only participant at Level 3 on the pre-survey remained at the same level post-survey.

When the participants' stock-flow thinking level was evaluated regarding majors, it was determined that two of the five mathematics PSTs were at Level 1, two were at Level 2, and one was at Level 3 on the pre-survey. In the post-survey, both Level 1 participants, one Level 2 participant, and one Level 3 participant did not improve and remained at the same level. It was also noted that one Level 2 participant unfortunately dropped back to Level 1. Interestingly, the pre-survey found that all nine science PSTs were at Level 2. Unfortunately, two participants dropped to Level 1 in the post-survey, and three remained at the same level, showing no improvement. Fortunately, four of them moved up to Level 3.

### *Causal Loop Thinking*

When Table 7 is analyzed in terms of levels of causal loop thinking from the pre-survey to the post-survey, the number of Level 2 respondents decreased, and the number of Level 3 respondents increased. Surprisingly, one participant at Level 0 on the pre-survey moved up to Level 3. Two of the four participants at Level 1 in the pre-survey moved to Level 2, and the other two moved to Level 3. Of the six participants at Level 2 on the pre-survey, one stayed the same, and the other five moved up to Level 3. Three participants at Level 3 on the pre-survey did not change levels on the post-survey.

When the participants' level of causal loop thinking was evaluated in terms of majors, it was found that one of the five mathematics PSTs was at Level 1, one was at Level 2, and the remaining three were at Level 3 in the pre-survey. In the post-survey, it is good to see that the Level 1 participant moved to Level 2 and the Level 2 participant moved to Level 3. The Level 3 participants did not progress and remained at the same level. In the pre-survey, one of the nine science PSTs was at Level 0, three were at Level 1, and the remaining five were at Level 2. In the post-survey, surprisingly, the Level 0 participant moved up to Level 3, one Level 1 participant to Level 2, and the remaining two Level 1 participants to Level 3. Four out of five Level 2 participants progressed to Level 3, but one remained at the same level and did not progress.

***Dynamic Thinking***

When Table 7 is analyzed in terms of levels of dynamic thinking from the pre-survey to the post-survey, it can be seen that the number of participants at Level 0 and Level 1 decreased, and the number of participants at Level 2 and Level 3 increased. Of the six participants at Level 0 on the pre-survey, two remained at the same level, two moved to Level 1, and the remaining two moved to Level 2 in the post-survey. All participants at Level 1 in the pre-survey moved to Level 2 in the post-survey. The participants at Level 2 in the pre-survey moved up to Level 3 in the post-survey.

When the dynamic thinking level of the participants was evaluated in terms of majors, it was found that three of the five mathematics PSTs were at Level 0, and the remaining two were at Level 1 in the pre-survey. In the post-survey, two participants at Level 0 moved up to Level 2, while one remained at the same level. It is good to see that all the participants at Level 1 in the pre-survey moved up to Level 2. In the pre-survey, three of the nine science PSTs were at Level 0, five were at Level 1, and the remaining one was at Level 2. In the post-survey, it can be seen that two of the three participants who were at Level 0 moved up to Level 2, and the remaining one remained at the same level with no improvement. It is nice to see that all of the participants at Level 1 in the pre-survey moved up to Level 2 in the post-survey. Only one participant at Level 2 on the pre-survey moved up to Level 3 on the post-survey.

## Discussion

A society with the potential to be systems literate and adept may be hampered in dealing with complex and dynamic situations by limited awareness and grasp of systems thinking. This study explored the change in STS among PSTs after participating in a systems thinking-oriented teacher education course. Jordan and colleagues (2013) demonstrated that instructional techniques based on the Structure-Behavior-Function Theory significantly improved understanding of the various factors involved, particularly concerning behaviors and functions. Similarly, improvements in the ability of pre-service teachers to answer scenario-based questions about structure, behavior, and purpose were observed in the current study.

Hmelo-Silver et al. (2007) identified challenges novice teachers face in understanding complex systems. PSTs focused on static system components, while in-service teachers focused on structural, functional, and behavioral elements (Hmelo-Silver & Pfeffer, 2004). Similarly, Lee et al. (2019) found that both in-service and PSTs struggled to identify system components, processes, and hidden dimensions. The present study's results align with these barriers, showing STS levels of PSTs before instruction were consistent with these difficulties. Their STS levels in the stock flow case before the course were slightly more advanced than in the causal loop and dynamic thinking cases, but still did not include novice-level data.

After the systems thinking course, the results from the developing level increased, and differences in skill levels were observed in the PSTs' performance in the scenario focusing on stock and flow thinking skills compared to the pre-survey. As mentioned in Aşık and Doğanca Küçük's study (2021), individuals' difficulty in understanding and solving stock-flow scenarios may be due to their decision-making processes rather than a lack of contextual knowledge about the tasks. Moreover, the majority of participants were unfamiliar with stock-flow scenarios, which could potentially impact their performance (Aşık & Doğanca Küçük, 2021). This situation underscores a key area for development in science education and suggests that strengthening teachers' decision-making abilities, especially in the context of complex systems, could greatly enhance their understanding and teaching of stock-flow concepts, leading to improved student learning outcomes in science classrooms (Karga & Ceyhan, 2024).

Perkins and Grotzer (2000) showed that when participants are asked to explain a collection of complicated systems, they frequently provide relatively basic causal explanations, as seen in the current study before instruction. Understanding the behavior and functions of a system requires a more detailed understanding of the underlying phenomena and their interrelationships. It was seen that after the intervention, the development of participants in causal loops and dynamic thinking cases was improved more explicitly. On the other hand, Davis et al. (2020) revealed that students who found more connections between variables performed better at identifying feedback. In the current study, although PSTs performed better in identifying causal interactions, they did not explicitly mention balancing or reinforcing causal loops as seen in the study conducted with science teachers (Karga & Ceyhan, 2024). In addition, another reason for the significant increase in participants' success in causal loops compared to the pre-survey may be that they worked on identifying stock-flow relationships before causal loops in the course timetable during the semester.

Ateskan and Lane (2018) found that after the workshop in the context of Education for Sustainable Development (ESD), teachers were more likely to see problems as a series of interconnected problems and that systems are constantly changing, which are the aspects of dynamic thinking. Also, Palmberg and colleagues (2017) showed that none of the PSTs acquired an intermediate or advanced level of systems thinking, incorporating interconnections, feedback, and behavioral components. Teachers struggle to deliver suitable learning experiences if they don't comprehend the nature of complex systems. In the study of Karaarslan Semiz and Teksöz (2020), science PSTs showed improvement in twelve ESD context aspects of systems thinking, including dynamic and cyclic thinking skills. Most of the teacher candidates had advanced to the developing or mastery level. Therefore, the findings of this study underscore the necessity for well-designed interventions designed to enhance the systems thinking skills of pre-service teachers across all three dimensions as emphasized by Yoon and colleagues (2017) the critical importance of professional development for teachers to effectively instruct on complex systems.

Although the current study used a context-independent measurement approach with scenarios, the PSTs showed improvement in dynamic thinking skills, as mentioned in the literature (Palmberg et al., 2017). However, the number of participants who still reach the highest level of dynamic thinking skills is very limited or low. This could be because interpreting gradual changes in a system, potential causal loops, and delays requires a deeper understanding and skills. Therefore, the initiatives to develop PSTs in the context of systems skills, such as those explored in this study, are important in preparing PSTs for the profession, as it was seen that fewer PSTs reached the upper level in terms of thinking with stock-flows and dynamic thinking skills. In addition, one of the interesting results of this study is that when the STS levels of mathematics and science PSTs are compared, the improvement in all three skills is more evident.

Moreover, the difference between the STS of mathematics and science PSTs may be due to less exposure to courses and learning materials that may contribute to the development of STS levels (Peretz et al., 2023). Considering the fluctuating systems thinking skill levels of pre-service teachers in three aspects, as indicated by the results, this finding suggests that improving and standardizing the duration and quality of both theoretical education and practical classroom experience in systems thinking may be beneficial. Such enhancements could lead to more consistent and developed systems thinking skills among teachers. On the other hand, the scenario-based questions employed in this study illuminate on the degree to which teachers' responses can be used to gauge students' systems thinking abilities. Addressing the shortcomings of assessments allows us to develop new techniques and improved instruments for assessing systems thinking abilities.

This exploratory research serves as a foundation for larger studies to validate and expand upon the results. While this study focuses on science and math PSTs, STS are relevant across various educational contexts and disciplines. Future research could include PSTs from other disciplines, such as social studies or language arts, to explore STS development in different domains. Future research could also involve longitudinal studies that follow the STS of PSTs as they progress through their teacher education programs and transition into their careers. By tracking participants' development over time and examining how they implement STS in real-world teaching scenarios, researchers can better understand the long-term impact of systems thinking-oriented interventions and identify potential barriers or facilitators to successfully integrating systems thinking into educational practice.

Furthermore, this study shows that involving PSTs in systems thinking is feasible and beneficial, even in the early stages of their education. Therefore, further research on systems thinking should inform curricula that integrate systems thinking into higher education programs (Elsawah et al., 2022; Karaaraslan Semiz & Teksöz, 2024). To gain a more comprehensive understanding of the impact of systems thinking-integrated courses across different educational levels, it is recommended to conduct studies with undergraduate and graduate programs.

In conclusion, this study contributes to the growing body of research on systems thinking in teacher education. However, future research should aim to build upon and extend its findings.  By pursuing the suggested directions, researchers can further advance our understanding of how to foster and assess STS among PSTs effectively. This will ultimately prepare PSTs to navigate the complexities of the modern educational landscape and equip their future students with the tools to thrive in an increasingly interconnected world.

## Compliance with Ethical Standards

*Disclosure of potential conflicts of interest*
The authors declare no competing interests.
*Funding*
No funds, grants, or other support was received for conducting this study.
*CRediT author statement*
Author 1: Conceptualization, Methodology, Formal Analysis, Investigation, Writing – original draft, Writing – review and editing, visualization, supervision
Author 2: Methodology, Formal Analysis, Investigation, Writing – original draft, Writing – review and editing, visualization
Author 3: Methodology, Formal Analysis, Investigation, Writing – original draft, Writing – review and editing
*Research involving Human Participants and/or Animals*
Approval to conduct the research was received from the Human Research Ethics Committee of Bogazici University (E-84391427-050.01.04-132885).

## Fen ve Matematik Öğretmen Adaylarının Sistem Düşüncesi Becerilerinin Vaka Senaryoları Aracılığıyla Değerlendirilmesi

**Özet:**

Karmaşık küresel sorunların ele alınması, sistem düşüncesi becerilerinin gerekliliğini vurgulayan daha kapsamlı ve bütüncül yaklaşımlar gerektirmektedir; ancak mevcut çalışmalar, öğretmen adaylarının sistem düşüncesi becerilerini anlamada önemli bir boşluk olduğunu göstermekte ve bu alanda daha fazla araştırma yapılması gerektiğini vurgulamaktadır. Bu keşifsel vaka çalışması araştırması, fen ve matematik öğretmen adaylarının sistem düşüncesi becerilerini senaryo tabanlı değerlendirmeler yoluyla araştırmıştır. Üç vaka senaryo örneği, sistem düşüncesinin belirli bir yönüne odaklanmıştır: stok akış, nedensel döngü ve dinamik düşünme. Bu çalışmanın katılımcıları, bir devlet araştırma üniversitesinin öğretmen eğitimi programında sistem düşüncesi dersi alan 14 öğretmen adayıdır. Veriler Sistem Düşüncesi Rubriği ve Dinamik Düşünme Becerileri Rubriği kullanılarak kodlanmıştır. Analizler, dinamik düşünme becerilerinde gelişmeler olduğunu göstermiş, ancak daha az sayıda katılımcı stok akışı düşünme konusunda gelişme göstermiştir. Fen ve matematik disiplinleri karşılaştırıldığında, fen bilgisi öğretmen adaylarının sistem odaklı düşünme becerilerinde matematik öğretmen adaylarına göre daha fazla ilerleme kaydettikleri görülmüştür. Bu ön araştırma, eğitimcilerde sistem düşüncesinin değerlendirilmesi ve geliştirilmesine yönelik içgörüler sunmaktadır. Sistem düşüncesi yaklaşımının öğretmen eğitimi programlarına entegre edilmesinin, öğretmenleri karmaşık sorunlarla etkili bir şekilde başa çıkmaya daha iyi hazırlayabileceğini öne sürmektedir. Öğretim yöntemlerinin kullanıldığı daha ileri çalışmalar, öğretmen adayları arasında sistem düşüncesi gelişiminin optimize edilmesine ışık tutabilir. Özünde bu araştırma, sistem düşüncesinin öğretmen eğitimini zenginleştirme potansiyelinin altını çizmektedir.

*Anahtar kelimeler*: Öğretmen adayları, sistem düşüncesi becerisinin ölçülmesi, sistem düşüncesi becerileri, öğretmen öğrenmesi.

## References

Arnold, R. D., & Wade, J. P. (2015). A definition of systems thinking: A systems approach. *Procedia Computer Science*, *44*, 669-678. https://doi.org/10.1016/j.procs.2015.03.050.

Aşık, G., & Doğança Kucuk, Z. (2021). Metacognition in action as a possible explanation for stock-flow failure. *System Dynamics Review*, *37*(4), 253-282. https://doi.org/10.1002/sdr.1692

Ateskan, A., & Lane, J. F. (2018). Assessing teachers' systems thinking skills during a professional development program in Turkey. *Journal of Cleaner Production, 172,* 4348-4356. https://doi.org/10.1016/j.jclepro.2017.05.094

Bartus, G. A., & Fisher, F. T. (2016). Barriers and opportunities to the acquisition of systems thinking skills for k-12 teachers. *Proceedings of ASME International Mechanical Engineering Congress and Exposition*, New York, 50571, V005T06A045. https://doi.org/10.1115/IMECE2016-67146

Batzri, O., Ben-Zvi Assaraf, O., Cohen, C. & Orion, N. (2015). Understanding the earth systems: Expressions of dynamic and cyclic thinking among university students. *Journal of Science Education and Technology*, *24*, 761-775. https://doi.org/10.1007/s10956-015-9562-8

Ben-Zvi Assaraf, O., & Orion, N. (2005). Development of system thinking skills in the context of earth system education. *Journal of Research in Science Teaching*, *42*(5), 518-560. https://doi.org/10.1002/tea.20061

Ben-Zvi Assaraf, O., & Orion, N. (2010). System thinking skills at the elementary school level. *Journal of Research in Science Teaching*, *47*(5), 540-563. https://doi.org/10.1002/tea.20351

Ben-Zvi Assaraf, O., Dodick, J., & Tripto, J. (2013). High school students' understanding of the human body system. *Research in Science Education*, *43*, 33-56. https://doi.org/10.1007/s11165-011-9245-2

Bielik, T., Delen, I., Krell, M., & Assaraf, O. B. Z. (2023). Characterising the literature on the teaching and learning of system thinking and complexity in STEM education: A bibliometric analysis and research synthesis. *Journal for STEM Education Research*, *6*(2), 1-33. https://doi.org/10.1007/s41979-023-00087-9

Boubonari, T., Papazoglou, D. N., Mogias, A., & Kevrekidis, T. (2023). Challenging Greek primary students' knowledge of ocean acidification using the carbon cycle context. *International Journal of Science and Mathematics Education*, *22*, 1-24. https://doi.org/10.1007/s10763-023-10431-5

Budak, U. S., & Ceyhan G. D. (2024). Research trends on systems thinking approach in science education, *International Journal of Science Education, 46*(5), 485-502. https://doi.org/10.1080/09500693.2023.2245106

Cooper, J. O., Heron, T. E., & Heward, W. L. (2019). *Applied behavior analysis* (3rd ed.). Pearson Education. https://books.google.com.tr/books/about/Applied_Behavior_Analysis.html?id=Yi4YygEACAAJ&redir_esc=y

Daniel, S., & Mazzurco, A. (2020). Development of a scenario-based instrument to assess co-design expertise in humanitarian engineering. *European Journal of Engineering Education*, *45*(5), 654-674. https://doi.org/10.1080/03043797.2019.1704689

Davis, K., Ghaffarzadegan, N., Grohs, J., Grote, D., Hosseinichimeh, N., Knight, D., Mahmoudi, H., & Triantis, K. (2020). The Lake Urmia vignette: A tool to assess understanding of complexity in socio-environmental systems. *System Dynamics Review*, *36*(2), 191-222. https://doi.org/10.1002/sdr.1659

Delaney, S., Ferguson, J. P., & Schultz, M. (2021). Exploring opportunities to incorporate systems thinking into secondary and tertiary chemistry education through practitioner perspectives. *International Journal of Science Education*, *43*(16), 2618-2639. https://doi.org/10.1080/09500693.2021.1980631

Doganca Kucuk, Z., & Saysel, A. K. (2018). Developing seventh grade students' understanding of complex environmental problems with systems tools and representations: a quasi-experimental study. *Research in Science Education*, *48*, 491-514. https://doi.org/10.1007/s11165-017-9620-8

Dorani, K., A. Mortazavi, A. Dehdarian, H. Mahmoudi, M. Khandan and A. N. Mashayekhi. (2015). Developing question sets to assess systems thinking skills, in *Proceedings of the 33rd International Conference of the System Dynamics Society*, Cambridge, MA, USA, (pp. 19-23). https://proceedings.systemdynamics.org/2015/papers/P1120.pdf

Eaton, A. C., Delaney, S., & Schultz, M. (2019). Situating sustainable development within secondary chemistry education via systems thinking: A depth study approach. *Journal* *of Chemical Education*, *96*(12), 2968-2974. https://doi.org/10.1021/acs.jchemed.9b00266

Eidin, E., Bielik, T., Touitou, I., Bowers, J., McIntyre, C., Damelin, D., & Krajcik, J. (2023). Correction to: Thinking in terms of change over time: Opportunities and challenges of using system dynamics models. *Journal of Science Education and Technology*, *22*, 1-28. https://doi.org/10.1007/s10956-023-10071-y

Elsawah, S., Ho, A. T. L., & Ryan, M. J. (2022). Teaching systems thinking in higher education. *INFORMS Transactions on Education*, *22*(2), 66-102. https://doi.org/10.1287/ited.2021.0248

Evagorou, M., Korfiatis, K., Nicolaou, C., & Constantinou, C. (2009). An investigation of the potential of interactive simulations for developing system thinking skills in elementary school: A case study with fifth-graders and sixth-graders. *International Journal of Science Education*, *31*(5), 655-674. https://doi.org/10.1080/09500690701749313

Fisher, D. M. & Systems Thinking Association. (2023). Systems thinking activities used in K-12 for up to two decades, *Frontiers in Education*, *8*, 1-14. https://doi.org/10.3389/feduc.2023.1059733

Forrester, J. W. (1994). System dynamics, systems thinking, and soft OR. *System Dynamics Review*. *10*(2–3), 245–256. https://doi.org/10.1002/sdr.4260100211

Forrester, J. (2007). System dynamics - A personal view of the first fifty years. *System Dynamics Review*, *23*, 245–358. https://doi.org/10.1002/sdr.382

Gay, L. R., Mills, G. E., & Airasian, P. W. (2012). *Educational research: Competencies for analysis and applications* (10th ed). Pearson. https://www.google.com.tr/books/edition/Educational_Research/J5eJVwEACAAJ?hl=tr

Hmelo-Silver, C. E., & Pfeffer, M. G. (2004). Comparing expert and novice understanding of a complex system from the perspective of structures, behaviors, and functions. *Cognitive Science*, *28*(1), 127-138. https://doi.org/10.1016/S0364-0213(03)00065-X

Hmelo-Silver, C. E., Marathe, S., & Liu, L. (2007). Fish swim, rocks sit, and lungs breathe: Expert-novice understanding of complex systems. *The Journal of the Learning Sciences, 16*(3), 307-331. https://doi.org/10.1080/10508400701413401

Hopper, M., & Stave, K. A. (2008). Assessing the effectiveness of systems thinking interventions in the classroom. In *The 26th International Conference of the System*

*Dynamics Society*, (pp. 1-26). Athens, Greece. https://proceedings.systemdynamics.org/2008/proceed/papers/STAVE390.pdf

Jordan, R. C., Hmelo-Silver, C., Liu, L., & Gray, S. A. (2013). Fostering reasoning about complex systems: Using the aquarium to teach systems thinking. *Applied Environmental Education & Communication*, *12*(1), 55-64. https://doi.org/10.1080/1533015x.2013.797860

Kali, Y., Orion, N., & Eylon, B. S. (2003). Effect of knowledge integration activities on students' perception of the Earth's crust as a cyclic system. *Journal of Research in Science Teaching*, *40*(6), 545-565. https://doi.org/10.1002/tea.10096

Karaarslan Semiz, G., & Teksöz, G. (2020). Developing the systems thinking skills of pre-service science teachers through an outdoor ESD course. *Journal of Adventure Education and Outdoor Learning*, *20*(4), 337-356. https://doi.org/10.1080/14729679.2019.1686038

Karaarslan Semiz, G., & Teksöz, G. (2024). Tracing system thinking skills in science curricula: A case study from Turkey. *International Journal of Science and Mathematics Education*, *22*(3), 515-536. https://doi.org/10.1007/s10763-023-10383-w

Karga, B., & Ceyhan, G. D. (2024). Investigating middle school science teachers' stock-flow, causal-loop, and dynamic thinking skills with scenario-based questions. *International Journal of Science Education*, 1-20. https://doi.org/10.1080/09500693.2024.2404546

Kriswandani, C. S. D., Irawati, S., & Hidayanto, E. (2022). Systemic thinking processes of preservice teachers with systematic cognitive style in solving complex problems. *Journal of Positive School Psychology*, 2706-2730. https://mail.journalppw.com/index.php/jpsp/article/view/6477

Lavi, R., & Dori, Y. J. (2019). Systems thinking of pre-and in-service science and engineering teachers. *International Journal of Science Education*, *41*(2), 248-279. https://doi.org/10.1080/09500693.2018.1548788

Lee, T. D., Gail Jones, M., & Chesnutt, K. (2019). Teaching systems thinking in the context of the water cycle. *Research in Science Education*, *49*, 137-172. https://doi.org/10.1007/s11165-017-9613-7

Mambrey, S., Timm, J., Landskron, J. J., & Schmiemann, P. (2020). The impact of system specifics on systems thinking. *Journal of Research in Science Teaching*, *57*(10), 1632-1651. https://doi.org/10.1002/tea.21649

Meadows, D. H. (2008). *Thinking in systems: A primer*. Chelsea Green Publishing. https://www.google.com.tr/books/edition/Thinking_in_Systems/CpbLAgAAQBAJ?hl=tr&gbpv=0

Mehren, R., Rempfler, A., Buchholz, J., Hartig, J., & Ulrich-Riedhammer, E. M. (2018). System competence modelling: Theoretical foundation and empirical validation of a model involving natural, social and human-environment systems. *Journal of Research in Science Teaching*, *55*(5), 685-711. https://doi.org/10.1002/tea.21436

Meilinda, M., Rustaman, N. F., Firman, H., & Tjasyono, B. (2018). Development and validation of the climate change system thinking instrument (CCSTI) for measuring system thinking on climate change content. *Journal of Physics: Conference Series, 1013*(1), 012046. https://doi.org/10.1088/1742-6596/1013/1/012046

Next Generation Science Standards (NGSS). (2013). *Next Generation Science Standards: For States*. https://www.nextgenscience.org/

Nuhoğlu, H. (2010). The effect of the system dynamics approach on understanding causal relationship skills in science education. *Procedia-Social and Behavioral Sciences, 2*(2), 3614-3618. https://doi.org/10.1016/j.sbspro.2010.03.561

Palmberg, I., Hofman-Bergholm, M., Jeronen, E., & Yli-Panula, E. (2017). Systems thinking for understanding sustainability? Nordic student teachers' views on the relationship between species identification, biodiversity and sustainable development. *Education Sciences*, *7*(3), 72. https://doi.org/10.3390/educsci7030072

Peretz, R., Tal, M., Akiri, E., Dori, D., & Dori, Y. J. (2023). Fostering engineering and science students' and teachers' systems thinking and conceptual modeling skills. *Instructional Science*, *51*(3), 509-543. https://doi.org/10.1007/s11251-023-09625-9

Perkins, D. N., & Grotzer, T. A. (2000). *Models and moves: Focusing on dimensions of causal complexity to achieve deeper scientific understanding*, (pp.24-28). Paper presented at the annual meeting of the American Educational Research Association, New Orleans, LA. https://eric.ed.gov/?id=ED441698

Plate, R. (2010). Assessing individuals' understanding of nonlinear causal structures in complex systems. *System Dynamics Review*, *26*(1), 19-33. https://doi.org/10.1002/sdr.432

Plate, R., & Monroe, M. (2014). A structure for assessing systems thinking. *The Creative Learning Exchange*, *23*(1), 1-3. https://doi.org/10.1002/sdr.432

Richmond, B. (1993). Systems thinking: critical thinking skills for the 1990s and beyond. *System Dynamics Review*, *9*(2), 113-133. https://doi.org/10.1002/sdr.4260090203

Riess, W., & Mischo, C. (2010). Promoting systems thinking through biology lessons. *International Journal of Science Education*, *32*(6), 705-725. https://doi.org/10.1080/09500690902769946

Rosenkränzer, F., Hörsch, C., Schuler, S., & Riess, W. (2017). Student teachers' pedagogical content knowledge for teaching systems thinking: Effects of different interventions. *International Journal of Science Education*, *39*(14), 1932–1951. https://doi.org/10.1080/09500693.2017.1362603

Salado, A., Chowdhury, A. H., & Norton, A. (2019). Systems thinking and mathematical problem solving. *School Science and Mathematics*, *119*(1), 49-58. https://doi.org/10.1111/ssm.12312

Senge, P. M. (1990). *The Fifth Discipline: The art and practice of the learning organization.* Crown. https://www.google.com.tr/books/edition/The_Fifth_Discipline/5jCLbnlmEuYC?hl=tr

Sommer, C., & Lücken, M. (2010). System competence – Are elementary students able to deal with a biological system?. *Nordic Studies in Science Education*, *6*, 125–143. https://doi.org/10.5617/nordina.255

Stave, K., & Hopper, M. (2007). What constitutes systems thinking? A proposed taxonomy. In *25th International Conference of the System Dynamics Society*. Boston, MA. http://static.clexchange.org/ftp/conference/CLE_2010/CO2010-06Session5MeasureST.pdf

Sweeney, L. B., & Sterman, J. D. (2000). Bathtub dynamics: Initial results of a systems thinking inventory. *System Dynamics Review*, *16*(4), 249-286. https://doi.org/10.1002/sdr.198

Sweeney, L., & Sterman, J. D. (2007). Thinking about systems: Student and teacher conceptions of natural and social systems. *System Dynamics Review*, *23*, 285–312. https://doi.org/10.1002/sdr.366

Taylor, S., Calvo-Amodio, J., & Well, J. (2020). A method for measuring systems thinking learning. *Systems*, *8*(2), 11. https://doi.org/10.3390/systems8020011

Tripto, J., Assaraf, O. B. Z., & Amit, M. (2018). Recurring patterns in the development of high school biology students' system thinking over time. *Instructional Science, 46*, 639-680. https://doi.org/10.1007/s11251-018-9447-3

Uskola, A., & Puig, B. (2022). Exploring primary preservice teachers' agency and systems thinking in the context of the COVID-19 pandemic. *Frontiers in Education,7*, p. 869643. https://doi.org/10.3389/feduc.2022.869643

Yoon, S. A., Anderson, E., Koehler-Yom, J., Evans, C., Park, M., Sheldon, J., Schoenfeld, I., Wendel, D., Scheintaub, H., & Klopfer, E. (2017). Teaching about complex systems is no simple matter: Building effective professional development for computer-supported complex systems instruction. *Instructional Science, 45*(1), 99–121. https://doi.org/10.1007/s11251-016-9388-7

Yin, R. K. (2009). *Case study research: Design and methods (4th ed.)*. Sage. https://www.google.com.tr/books/edition/Case_Study_Research/FzawIAdilHkC?hl=tr&gbpv=0

York, S., Lavi, R., Dori, Y. J., & Orgill, M. (2019). Applications of systems thinking in STEM education. *Journal of Chemical Education*, *96*(12), 2742-2751. https://doi.org/10.1021/acs.jchemed.9b00261

Necatibey Faculty of Education Electronic Journal of Science and Mathematics Education
Vol. 18, Issue 2, December 2024, pp. 404-429. ISSN: 1307-6086

Research Article

# Uncovering the Undergraduate Physiotherapy Students' Conceptual Understanding Levels and Misconceptions of Simple Electric Circuits from 2018 to 2023

**Kübra ÖZMEN[1]**

[1] Başkent University, Faculty of Health Sciences, Türkiye, kubraozmen2017@gmail.com, http://orcid.org/0000-0001-7838-8314

Received : 09.08.2024 Accepted : 05.11.2024

Doi: https://doi.org/10.17522/balikesirnef.1530869

---

*Abstract* – This study aimed to investigate undergraduate physiotherapy students' conceptual understanding levels of simple electric circuits and the misconceptions they brought to introductory physics courses, over five years, from 2018 to 2023. This retrospective study adopted an exploratory research methodology. Using purposive sampling, undergraduate students majoring in physiotherapy and rehabilitation and attending the Physics II course at a private university from 2018 to 2023 were selected as participants. In total, the results of the Simple Electric Circuit Diagnostic Test (SECDT) voluntarily taken by 296 students (209 females and 87 males) aged 18 to 22 were included in the study. The students' responses in the SECDT were analyzed using MS Excel and further analysis with IBM PASW18. The researcher calculated the frequency of each misconception category by adding up the students' responses to all questions related to that category. The independent samples Kruskal Wallis test with post-hoc analysis (pairwise comparisons) was conducted to determine whether student groups in different academic years were statistically different regarding SECDT scores. The results showed that students held five common misconceptions over five years: clashing current model, short circuit misconception, power supply as constant current source model, sequential reasoning, and local reasoning. Moreover, the study revealed a peak in 2021-2022 regarding students having better conceptual understanding levels than other years. Five-year data showed that minimal changes have occurred, and that prevalent misconceptions did not change over time. These findings can help educators when designing their courses, explicitly focusing on these misconceptions to promote a better conceptual understanding of electricity.

*Keywords*: Physics for non-physicists, non-physics science majors, conceptual understanding levels, health sciences, misconception, physiotherapy, simple electric circuits.

------------

Corresponding author: Kübra ÖZMEN, kubraozmen2017@gmail.com

## Introduction

Physics is often perceived as abstract and disconnected from students' everyday experiences (Kollöffel & de Jong, 2013; Ramma et al., 2017). Research in physics education has revealed that many introductory physics students struggle to grasp physics concepts (Neidorf et al., 2020). Developing these concepts is intricate and influenced by individuals' perceptions, experiences, and information acquisition habits. Life experiences differ from person to person, resulting in varying interpretations of the same subject. Language used in the classroom, textbooks, and even teachers can be primary sources of misconceptions (Hammer, 1996). It is crucial for instructors to identify and address these misconceptions to foster student engagement and facilitate a better grasp of scientific principles (Uwamahoro et al., 2021).

One abstract concept that deserves attention is electricity. Although invisible, it is omnipresent and essential to numerous crucial processes, such as powering electronic devices and enabling medical equipment. Understanding electricity is fundamental across all educational levels in physics, as it establishes a strong foundation for comprehending more complex subjects (Mbonyiryivuze et al., 2019; Mulhall et al., 2001). Studies in science education have focused on grasping conceptual understanding, particularly in simple electric circuits, both in practical and theoretical contexts (Ünal, 2022). Earlier studies have identified misconceptions about electricity (Chambers & Andre, 1997; Engelhardt & Beichner, 2004; Peşman & Eryılmaz, 2010; Shipstone, 1984).

It is important for students to have a good understanding of electricity, as it plays a significant role in various fields such as engineering, technology, and medicine (Melzer et al., 2012). The fundamental principles of electric circuits are crucial, as they form the basis for students' comprehension of technology (Önder et al., 2017). In this regard, physics courses designed for health science students should aim to provide essential conceptual understanding, building on their existing knowledge of physics, as argued by Özmen (2024). Physics is a fundamental science relevant to numerous professions (Ozdemir et al., 2020). Professionals such as physiotherapists rely on a foundational understanding of electricity (Espejo-Antúnez et al., 2022; Kollöffel & de Jong, 2013; Martellucci, 2015). This knowledge helps them use electrical modalities safely and effectively, comprehend the bioelectrical functions of the human body, keep up with technological advancements, and collaborate efficiently in healthcare settings (Melzer et al., 2012). Ultimately, this understanding contributes to improved patient care and treatment outcomes. Accordingly, this study aimed to investigate the misconceptions undergraduate physiotherapy students brought to introductory physics courses over five years, from 2018 to 2023.

This paper focuses on the conceptual understanding levels of physiotherapy and rehabilitation (PR) students related to simple electric circuits and investigates the following research questions:

i. How did PR students' conceptual understanding levels of simple electric circuits change between 2018 and 2023?
ii. Which misconceptions about simple electric circuits persisted among PR students between 2018 and 2023?
iii. Were there any significant differences between PR students' conceptual understanding levels and misconceptions of simple electric circuits over five years?

## Background of the Study

Understanding key electrical concepts and the distinction between short and open circuits is essential for grasping the workings of simple electric circuits. The literature contains ample evidence of misconceptions about simple electrical circuits, common even among students who have received formal instruction in the relevant material (Quezada-Espinoza et al., 2023). Many students often deem electric circuits complex and abstract, with several studies highlighting their learning challenges (Taslidere & Yıldırım, 2023). Consequently, extensive research has identified misconceptions surrounding simple electric circuits (Chambers & Andre, 1997; Peşman & Eryılmaz, 2010; Turgut et al., 2011). Early research on the subject demonstrated that students struggle with the concept of electricity (Chambers & Andre, 1997; Shipstone, 1984).

Previous research has extensively documented several common misunderstandings about simple electric circuits. For instance, students often incorrectly assume that electric current is used up as it travels through a circuit, which can hinder learning about current conservation and the Kirchhoff's laws (Shipstone, 1984). Another prevalent misconception is that batteries provide a constant current rather than a constant voltage, which can confuse students when analyzing circuit behavior under various loads (McDermott & Shaffer, 1992). Confusion about the direction of current flow, particularly the difference between conventional current (positive to negative) and electron flow (negative to positive), is also widespread (Dupin & Johsua, 1987).

Students often misunderstand the role of a voltage source, thinking that it pushes the current through a circuit without realizing that it actually establishes a potential difference driving the current (Fredette & Lochhead, 1980). Many students incorrectly assume that the voltage is the same at all points in a circuit, failing to recognize that the potential difference between two points which drives the current flow. Misconceptions about series and parallel circuits are also common, especially regarding the voltage and the current distribution.

Students frequently fail to understand that the voltage across each branch is the same in parallel circuits, while in series circuits, the current through each component is identical (Cohen et al., 1983). In parallel circuits, they may fail to understand how the current is divided among branches, affecting each branch's total resistance and their individual currents (Chambers & Andre, 1997). Researchers have also reported specific misconceptions about electricity and electric circuits in the literature (Engelhardt & Beichner, 2004; Peşman & Eryılmaz, 2010).

In their review, Duman and Avcı (2014) examined the studies that identified the misconceptions of secondary school students and pre-service science and technology teachers in science education between 2003 and 2013 in Türkiye-based university journals available electronically. They detected two studies related to misconceptions: about simple electric circuits (Ateş & Polat, 2005) and about the electric current (Yıldırım et al., 2008). In a similar study, Yeltekin-Atar et al. (2021) examined the studies on misconception and conceptual understanding in physics education published between 2010 and 2019 within all volumes of the faculty of education journals. The results showed that "electricity and magnetism" is the third most common physics content in misconception and conceptual understanding studies (in 13 out of 90 studies).

However, previous studies differ in their prevalence of common misconceptions related to simple electric circuits. Ateş and Polat (2005) identified the following misconceptions in their study with university students: sequential reasoning (41%), the power supply as a constant current source model (38%), the shared current model (27%), the attenuation model (7%), the unipolar model (4%), and the clashing current model (5%). Another study by Yıldırım et al. (2008) focused on 6-8$^{th}$-grade students and their understanding of electricity. It revealed that most students had misconceptions about electric currents, consistently across all three grades, and that students often struggled with basic and abstract concepts of current and voltage. Additionally, when changes were made to electric circuits, such as adding a resistance, students had difficulty in distinguishing between equivalent current and voltage concepts. They also faced challenges in grasping the effects of resistance on current and voltage, leading to misconceptions.

Looking at the studies in the literature, Çepni and Keleş (2006) listed the most common misconceptions as: the power-consumption model, the clashing currents model, and the attenuation model. In a study conducted with pre-service science teachers, Altun (2009) reported that short circuit, shared current, attenuation, and clashing currents models are the most common misconceptions. Peşman and Eryılmaz (2010) found the shared current model, the clashing current model, short circuit misconception, power supply as a constant current source, and the local reasoning model as the most common misconceptions.

In Karakuyu and Tüysüz's (2011) study involving tenth-grade students, one common misconception was the belief that adding or removing a resistance from the circuit would always increase or decrease the total resistance, regardless of how the resistors were connected.

Another extensive study conducted by Eryılmaz et al. (2016) aimed to assess misconceptions across different regions in Türkiye by administering the Simple Electric Circuits Diagnostic Test (Peşman, 2005) to a large sample of high school students (N=11550). The study revealed that the most prevalent misconception (31%) among students regarding simple electrical circuits was the clashing current model.

In a study conducted with students in physiotherapy and rehabilitation major, the clashing current and the local reasoning models were found to be more common misconceptions (Özmen, 2019). Similarly, in Özmen's (2022) study, the most common misconceptions among the students in audiology major were the clashing current and the local reasoning models, where the students consider the electric circuit locally ignoring that the events in the whole circuit coincide.

These misconceptions can significantly impact learning, leading to difficulties in understanding more advanced topics and applying theoretical knowledge in practical situations (Reiner et al., 2000). Misconceptions can persist despite instruction, suggesting that traditional teaching methods may be ineffective in addressing these issues (Özmen, 2024; Vosniadou, 1994). Previous research has shown that misconceptions about electric circuits are resistant to change because they are often based on intuitive but incorrect interpretations of physical phenomena (Reiner et al., 2000). Therefore, it is crucial for educators to identify these misconceptions early and address them through targeted instructional interventions.

Educators can develop more effective teaching strategies that enhance student learning and comprehension in introductory physics courses by understanding the most frequent misconceptions related to simple electric circuits.

### *Assessment of Misconceptions with Multiple Choice Diagnostic Tests*

Multiple-choice tests are commonly used in educational assessments as they evaluate the knowledge of the students and pinpoint misunderstandings (Kaltakci-Gurel et al., 2015). To address these misunderstandings, researchers have developed one, two, and three-tier multiple-choice tests, each providing varying levels of diagnostic precision. One-tier multiple-choice tests are the most conventional type, consisting of a single question and several answer choices. These tests are efficient and straightforward to administer, making them popular in large classroom settings. However, one-tier tests often do not fully uncover the reasoning behind students' choices. While they can identify incorrect answers, they do not offer insight into the misconceptions that led to those answers (Haladyna, 2004).

Two-tier multiple-choice tests were created to tackle this limitation. The first tier comprises a traditional multiple-choice question, while the second tier requires students to explain their choice. This format enables educators to gain insight into the students' thought processes and pinpoint specific misunderstandings. For example, when assessing understanding of Ohm's Law, the first tier might inquire about the current through a resistor given a certain voltage and resistance, while the second tier would request the student to explain their reasoning (Treagust, 1988). This additional layer of questioning helps differentiate between different types of misunderstandings and provides more detailed diagnostic information.

Three-tier multiple-choice tests add another layer of depth by incorporating a confidence rating. After answering the initial question and providing a reason, students indicate their confidence in their answers. This tier helps to identify not only the presence of misunderstandings but also the students' certainty in their incorrect beliefs (Caleon & Subramaniam, 2010; Peşman & Eryılmaz, 2010). For instance, in a three-tier test on electric circuits, a student might select an incorrect answer about parallel circuit behavior, provide a flawed explanation, and then rate their confidence as high. This combination of responses indicates a strong and potentially resistant misconception that requires targeted instructional intervention.

# Method

## Research Design

This retrospective study used exploratory research methodology in a longitudinal manner. Longitudinal studies involve collecting continuous or repeated data to track specific individuals over long periods (Caruana et al., 2015). These studies are typically observational, as researchers gather quantitative and/or qualitative data on various exposures and outcomes without introducing any external influence (Rezigalla, 2020). Longitudinal research can take various forms and this study adopted a "repeated cross-sectional" research design, where participants differed entirely on each sampling occasion.

## Participants

Using purposive sampling, the study retrospectively analyzed the results of the Simple Electric Circuit Diagnostic Test (SECDT) voluntarily taken by undergraduate students in physiotherapy and rehabilitation major enrolled in the Physics II course at Başkent University from 2018 to 2023. In total, 296 students were included in the study, with 209 females and 87 males. Table 1 presents the distribution of participants by gender over the five-year period.

**Table 1** The Gender Distribution of the Participants across Five Years

| Year | Female N (%) | Male N (%) | Total N (%) |
|---|---|---|---|
| 2018-2019 | 50 (84.7) | 9 (15.3) | 59 (19.9) |
| 2019-2020 | 39 (72.2) | 15 (27.8) | 54 (18.2) |
| 2021-2022 | 55 (82.1) | 12 (17.9) | 67 (22.6) |
| 2022-2023 | 23 (60.5) | 15 (39.5) | 38 (12.8) |
| 2023-2024 | 42 (53.8) | 36 (46.2) | 78 (26.4) |
| Total | 209 (70.6) | 87 (29.4) | 296 (100) |

## Instrument

The Simple Electric Circuit Diagnostic Test (SECDT) was utilized in this study to evaluate students' conceptual understanding level. The SECDT is a three-tier diagnostic test developed by Peşman (2005) that measures 11 misconception models related to simple electric circuits through 12 questions (p. 157). The first set of questions consists of multiple-choice queries with typically two or three options. The subsequent set requires justifications for the chosen answer from the first tier, and the final set evaluates the students' confidence in their responses to the previous questions. By combining answers from the first and second tiers and responding to "sure" in the third tier, specific misconception models based on the answers are represented. These models include the sink model (M1), the attenuation model

(M2), the shared–current model (M3), the clashing current model (M4), the empirical rule model (M5), the short circuit model (M6), the power supply as a constant current source model (M7), the parallel circuit model (M8), the sequential reasoning model (M9), the local reasoning model (M10), and the current flow as water flow model (M11). For a comprehensive understanding of the SECDT structure, it is advisable to refer to Table 1 in Peşman and Eryılmaz’s (2010) study (p. 212). The reliability coefficient of the test provided in the original research was calculated to be 0.33 for misconception scores and 0.69 for correct scores using the Cronbach's alpha. For the current study, the reliability coefficient is calculated as 0.41 for misconception scores; and as 0.70 for correct scores. The reliability coefficients computed from the misconceptions scores are typically lower than those calculated from the correct scores (Kaltakci-Gurel et al., 2017). The reliability coefficients in this study are similar to those in Peşman (2005) , representing the nature of misconception tests (Arı et al., 2017).

**Data Collection**

The study was approved by the Başkent University Institutional Review Board and Ethics Committee (Project no: KA23/284) and supported by the Başkent University Research Fund. Data collection occurred yearly at the beginning of the spring semester during the second week of the Physics II course. Students were informed about the study during the first week of the course, and additional (bonus) points were given to students who participated voluntarily. Data collection was completed within a 50-minute lecture. Data was not collected during the 2020-2021 spring semester because the author was on maternal leave. Additionally, due to February 2023 Kahramanmaraş earthquakes, the data collection was conducted online during the 2022-2023 spring semester. A low response rate was noted due to the Council of Higher Education’s (CoHE, 2023) declaration that made the participation in online courses non-mandatory.

**Data Analysis**

The students' responses in the SECDT were analyzed using the MS Excel and the IBM PASW18. The researcher calculated the frequency and percentage of each misconception category by adding up students' responses to all questions related to that category. Any misconception with a rate of 10% or higher is regarded as a serious misconception in this study, as Caleon and Subramaniam (2010) recommended. Since the assumption of normal distribution for the dependent variables was violated, the independent samples Kruskal Wallis test with post-hoc analysis (pairwise comparisons), a non-parametric test, was conducted to determine whether student groups (clusters) in different academic years were statistically different in terms of SECDT scores.

## Results

### Descriptive Results of PR Students' Responses to the SECDT

The descriptive statistics of SECDT across different academic years indicate variations in students' performance (Figure 1). The mean scores varied, with the lowest in 2019-2020 at 3.76 and the highest in 2021-2022 at 5.91. The maximum score varies slightly between 9 and 10 within years, except the peak at 12 in 2021-2022. The data suggests varying student performance levels across different years, with notable improvements in certain periods and declines in others.

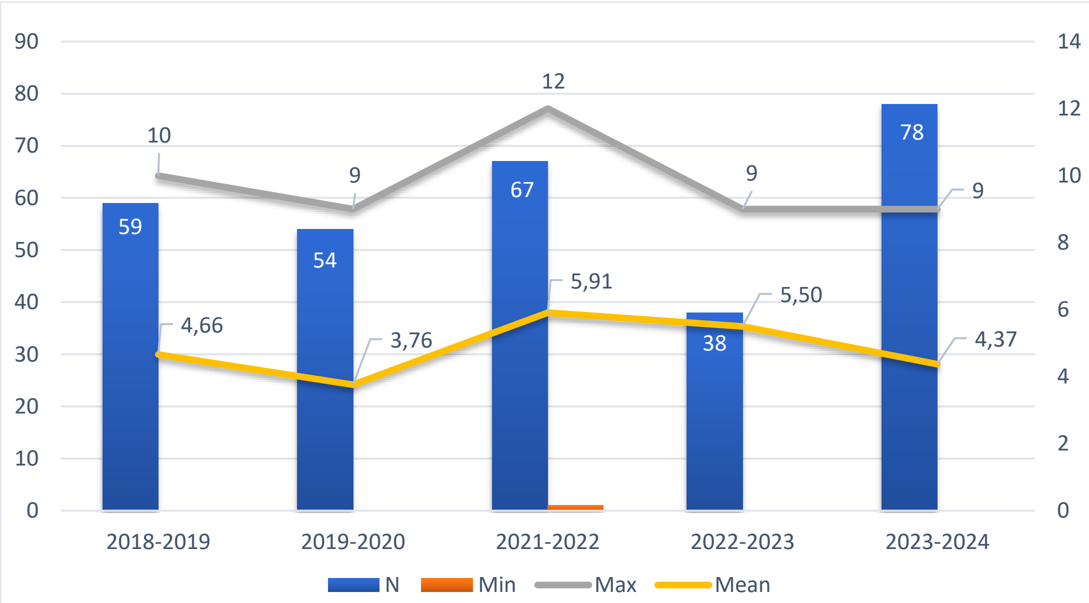



**Figure 1** Descriptive Statistics of the SECDT scores from 2018 to 2023

Table 2 outlines the percentage of correct responses students gave in the SECDT. Out of 12 questions, students had a correct response rate of less than 30% in five questions (Q2, Q3, Q5, Q9, and Q10).

**Table 2** Percentage of Students' Correct Responses for Each Question in SECDT

| Question No. | Correct responses - percentage (%) 2018-2019 | 2019-2020 | 2021-2022 | 2022-2023 | 2023-2024 |
|---|---|---|---|---|---|
| Q1 | 42.4 | 42.6 | **70.1** | 44.7 | 39.7 |
| Q2 | **8.5** | 7.4 | 1.5 | 5.3 | 5.1 |
| Q3 | **15.3** | 14.8 | 10.4 | 13.2 | 9.0 |
| Q4 | 50.8 | 42.6 | 82.1 | **86.8** | 67.9 |
| Q5 | 5.1 | 0.0 | 6.0 | **7.9** | 2.6 |
| Q6 | 59.3 | 51.9 | **86.6** | 76.3 | 65.4 |
| Q7 | 64.4 | 53.7 | 77.6 | **84.2** | 61.5 |
| Q8 | **59.3** | 38.9 | 58.2 | 44.7 | 43.6 |
| Q9 | 23.7 | 16.7 | 26.9 | **28.9** | 12.8 |
| Q10 | 10.2 | 7.4 | 13.4 | **15.8** | 14.1 |
| Q11 | 67.8 | 64.8 | **85.1** | 81.6 | 67.9 |
| Q12 | 61.0 | 35.2 | **73.1** | 60.5 | 52.6 |

*Boldface numbers refer to the highest percentages within year clusters.

Q2 analyzes a parallel circuit to determine the current flowing through each branch. Q3 examines how the magnitude of electric current changes in a series of bulb connections. Similarly, Q5 requires students to compare the magnitude of electric current when a new bulb is connected in parallel to a simple electric circuit. In Q9, students are asked to compare the brightness of a bulb connected with two identical resistors in series (10$\Omega$-bulb-10$\Omega$). This is done by increasing the resistance of one resistor in the electric circuit (10$\Omega$-bulb-20$\Omega$) and changing the location of resistors (20$\Omega$-bulb-10$\Omega$). Lastly, Q10 inquires whether a bulb gives light if one point of the bulb touches the current-carrying wire.

## Distribution of Misconception Models Related to Simple Electric Circuits

The distribution of the sink model misconception (M1) across different academic years is detected by three alternative responses given to Q1 and Q10 (M1_1: 1.1 a, 1.2 a, 1.3 a; M1_2: 10.1 a, 10.2 b, 10.3 a; M1_3: 10.1 b, 10.2 b, 10.3 a). Q1 asks whether a lamp will give light if it is connected to the positive terminal of a battery with a connecting wire. On the contrary, the same question is asked in Q10 when a wire is connected to a battery's negative and positive terminals while the lamp touches the wire. If a student believes that a lamp will light up when connected to the positive terminal of a battery with a wire, they might have the sink model misconception. Similarly, if a student confidently explains that the lamp will or will not light up because "the bulb is connected to the positive terminal," they may also have the sink model. Occurrences of the misconception were relatively low. However, the frequency and percentage of M1_3 increased considerably over time, reaching 14.1% in

2023-2024. The data underscores the persistence and increase of specific misconceptions about the sink model among students over the years (see Appendix).

The attenuation model misconception (M2) across different academic years was measured with two alternative responses to Q4 in the SECDT (M2_1: 4.1 c, 4.2 c, 4.3 a; M2_2: 4.1 b, 4.2 c, 4.3 a.). In Q4, two identical bulbs are connected with a battery in series, and the magnitudes of currents in each connection and the brightness of the bulbs are compared. If the students confidently explain that the magnitude of electric current will decay as the bulbs consume it, they may hold the attenuation model misconception. M2_1 shows a gradual increase in occurrences, from 1.7% in 2018-2019 to 7.7% in 2023-2024, indicating a rising trend in this alternative response over the years (see Appendix). However, the findings suggest that this misconception has either been non-existent or has not been effectively addressed recently.

The shared current model misconception (M3) across different academic years was measured by five alternative responses to Q3, Q4, and Q5 (M3_1: 3.1 b, 3.2 c, 3.3.a; M3_2: 3.1 a, 3.2 c, 3.3 a; M3_3: 4.1 d, 4.2 c, 4.3 a; M3_4: 5.1 b, 5.2 c, 5.3 a; M3_5: 5.1 a, 5.2 c, 5.3 a). In these questions, students have a misconception that electrical devices equally share an electrical current independent of how they are connected. The occurrence of the misconception varied slightly over the years, especially for M3_1, M3_2, and M3_5 (see Appendix). For example, M3_2 shows notable fluctuations, peaking at 10.5% in 2022-2023.

The distribution of the clashing current model misconception (M4) across different academic years is presented in Table 3. This model is detected by two alternative responses to Q1 and Q10 (M4_1: 1.1 b, 1.2 b, 1.3 a; M4_2: 10.1 a, 10.2 a, 10.3 a). If students respond to these questions that positive and negative charges should meet in the bulb to give light, they may have a clashing current model. M4_1 shows a marked increase over the years, beginning at 27.1% in 2018-2019 and rising to 48.7% in 2023-2024, indicating a growing prevalence of this misconception among students. Similarly, M4_2 began at 10.2% in 2018-2019, dropped to 3.7% in 2019-2020, then surged to 11.9% in 2021-2022, peaking at 21.1% in 2022-2023 before slightly declining to 7.7% in 2023-2024. These trends highlight a persistent challenge with the clashing current model misconception, with particularly high occurrences of M4_1.

**Table 3** Distribution of Clashing Current Model Misconception (M4)

| Year | M4_1 | | M4_2 | |
|---|---|---|---|---|
| | f | % | f | % |
| 2018-2019 | 16 | 27.1 | 6 | 10.2 |
| 2019-2020 | 13 | 24.1 | 2 | 3.7 |
| 2021-2022 | 16 | 23.9 | 8 | 11.9 |
| 2022-2023 | 18 | 47.4 | 8 | 21.1 |
| 2023-2024 | 38 | 48.7 | 6 | 7.7 |

*M4_1: 1.1 b, 1.2 b, 1.3 a; M4_2: 10.1 a, 10.2 a, 10.3 a.

The empirical rule model misconception (M5) was assessed through three different responses to Q4, Q7, and Q12 (M5_1: 4.1 b, 4.2 a, 4.3 a; M5_2: 7.1 b, 7.2 b, 7.3 a; M5_3: 12.1 a, 12.2 b, 12.3 a). The questions inquire how students perceive the brightness of bulbs when they are connected in series (Q4), connected in parallel (Q7), and connected in parallel with a conductive wire (Q12). Students who hold empirical rule model respond to these questions as the bulb will shine brighter if it is positioned near the battery in a circuit. The responses for M5_1, M5_2, and M5_3 consistently remained low (see Appendix). Overall, it seems that empirical model misconceptions are less frequent compared to other misconceptions.

The distribution of misconceptions about short circuits (M6) among students in different academic years is presented in Table 4. The model is identified through four responses to questions 8, 10, and 12 (M6_1: 8.1 b, 8.2 b, 8.3 a; M6_2: 8.1 c, 8.2 c, 8.3 a; M6_3: 10.1 a, 10.2 c, 10.3 a; M6_4: 12.1 b, 12.2 d, 12.3 a). Students who hold short circuit misconception responds to these questions without recognizing the conductive wire connections in an electric circuit. M6_1 is the most commonly chosen alternative, with occurrences rising significantly from 11.9% in 2018-2019 to a peak of 19.2% in 2023-2024. M6_2 remains consistently low with minimal variation. On the other hand, M6_3 showed fluctuations, reaching a notable high of 15.8% in 2022-2023. Similarly, M6_4 also displayed variations, peaking at 9.0% in 2023-2024. These trends indicate that while M6_2 is less common, M6_1 and M6_3 are more frequently observed, signaling ongoing challenges in students' comprehension of the short circuit concept.

**Table 4** Distribution of Short Circuit Misconception (M6)

| Year | M6_1 | | M6_2 | | M6_3 | | M6_4 | |
|---|---|---|---|---|---|---|---|---|
| | f | % | f | % | f | % | f | % |
| 2018-2019 | 7 | 11.9 | 2 | 3.4 | 1 | 1.7 | 3 | 5.1 |
| 2019-2020 | 10 | 18.5 | 2 | 3.7 | 3 | 5.6 | 2 | 3.7 |
| 2021-2022 | 9 | 13.4 | 3 | 4.5 | 7 | 10.4 | 1 | 1.5 |
| 2022-2023 | 3 | 7.9 | 4 | 10.5 | 6 | 15.8 | 3 | 7.9 |
| 2023-2024 | 15 | 19.2 | 2 | 2.6 | 10 | 12.8 | 7 | 9.0 |

*M6_1: 8.1 b, 8.2 b, 8.3 a; M6_2: 8.1 c, 8.2 c, 8.3 a; M6_3: 10.1 a, 10.2 c, 10.3 a; M6_4: 12.1 b, 12.2 d, 12.3 a.

Table 5 provides an analysis of the distribution of the power supply as constant current misconception (M7) across five academic years by four alternative responses to Q3, Q5, and Q9 (M7_1: 3.1 c, 3.2 a, 3.3 a; M7_2: 3.1 a, 3.2 a, 3.3 a; M7_3: 5.1 c, 5.2 e, 5.3 a; M7_4: 9.1 d, 9.2 d, 9.3 a). The data reveals a varying prevalence of these misconceptions over the years. M7_1 gradually increased from 11.9% in 2018-2019 to 32.1% in 2023-2024, peaking at 34.3% in 2021-2022. M7_2 and M7_4 responses were consistently low. M7_3 showed a significant rise, starting from 37.3% in 2018-2019, peaking dramatically at 77.6% in 2021-2022, before slightly declining to 50.0% in 2023-2024. Overall, the data highlights persistent and fluctuating trends in misconceptions about constant current in power supplies, with M7_1 and M7_3 being notably prevalent, while M7_2 and M7_4 remained relatively uncommon.

**Table 5** Distribution of Power Supply as Constant Current Misconception (M7)

| Year | M7_1 | | M7_2 | | M7_3 | | M7_4 | |
|---|---|---|---|---|---|---|---|---|
| | f | % | f | % | f | % | f | % |
| 2018-2019 | 7 | 11.9 | 0 | 0.0 | 22 | 37.3 | 0 | 0.0 |
| 2019-2020 | 6 | 11.1 | 0 | 0.0 | 12 | 22.2 | 0 | 0.0 |
| 2021-2022 | 23 | 34.3 | 1 | 1.5 | 52 | 77.6 | 2 | 3.0 |
| 2022-2023 | 13 | 34.2 | 0 | 0.0 | 21 | 55.3 | 0 | 0.0 |
| 2023-2024 | 25 | 32.1 | 1 | 1.3 | 39 | 50.0 | 5 | 6.4 |

*M7_1: 3.1 c, 3.2 a, 3.3 a; M7_2: 3.1 a, 3.2 a, 3.3 a; M7_3: 5.1 c, 5.2 e, 5.3 a; M7_4: 9.1 d, 9.1 d, 9.3 a.

One response alternative to Q5 in SECDT measured the parallel circuit misconception (M8) distribution (M8_1: 5.1 a, 5.2 a, 5.3 a.). The findings indicated that this misconception was not very common among students (see Appendix).

Table 6 displays the distribution of misconceptions related to sequential reasoning (M9) over five academic years, considering two alternative responses to Q9 (M9_1: 9.1 a, 9.2 a, 9.3 a; M9_2: 9.1 c, 9.2 b, 9.3 a). The data reveals consistent and fluctuating patterns in misconceptions about sequential reasoning with M9_1, while M9_2 remained relatively uncommon.

**Table 6** Distribution of Sequential Reasoning Misconception (M9)

| Year | M9_1 | | M9_2 | |
|---|---|---|---|---|
| | f | % | f | % |
| 2018-2019 | 4 | 6.8 | 1 | 1.7 |
| 2019-2020 | 7 | 13.0 | 0 | 0.0 |
| 2021-2022 | 16 | 23.9 | 0 | 0.0 |
| 2022-2023 | 6 | 15.8 | 6 | 15.8 |
| 2023-2024 | 15 | 19.2 | 2 | 2.6 |

*M9_1: 9.1 a, 9.2 a, 9.3 a; M9_2: 9.1 c, 9.2 b, 9.3 a.

In Table 7, the distribution of local reasoning misconception (M10) is displayed across different academic years by three alternative responses to Q2, Q5, and Q12 (M10_1: 2.1 a, 2.2 a, 2.3 a; M10_2: 5.1 a, 5.2 b, 5.3 a; M10_3: 12.1 a, 12.2 c, 12.3 a). The response to M10_1 is quite prevalent throughout the years, peaking at 85.1% in 2021-2022. However, the prevalence of responses to Q5 and Q12 are quite low when compared to the M10_1 response given for Q2.

**Table 7** Distribution of Local Reasoning Misconception (M10)

| Year | M10_1 | | M10_2 | | M10_3 | |
|---|---|---|---|---|---|---|
| | f | % | f | % | f | % |
| 2018-2019 | 29 | 49.2 | 1 | 1.7 | 0 | 0.0 |
| 2019-2020 | 32 | 59.3 | 4 | 7.4 | 2 | 3.7 |
| 2021-2022 | 57 | 85.1 | 2 | 3.0 | 4 | 6.0 |
| 2022-2023 | 30 | 78.9 | 1 | 2.6 | 0 | 0.0 |
| 2023-2024 | 60 | 76.9 | 3 | 3.8 | 2 | 2.6 |

*M10_1: 2.1 a, 2.2 a, 2.3 a; M10_2: 5.1 a, 5.2 b, 5.3 a; M10_3: 12.1 a, 12.2 c, 12.3 a

Lastly, the current flow as water flow misconception model (M11) was measured by three alternative responses to Q6, Q7, and Q11 (M11_1: 6.1 a, 6.2 a, 6.3 a; M11_2: 7.1 c, 7.2 a, 7.3 a; M11_3: 11.1 a, 11.2 b, 11.3 a). The analysis revealed that this misconception was not very common among students (see Appendix).

To sum up, over five years, the results indicated that students held five common misconceptions that they might bring from K-12 education and/or their daily life experiences: the clashing current model (M4), the short circuit misconception model (M6), the power supply as constant current source model (M7), the sequential reasoning (M9), and the local reasoning model (M10).

## Comparison of PR Students' Conceptual Understanding Levels of Simple Electric Circuits Across Five Years

The student responses to the SECDT were coded as correct (i.e., 1st tier: correct, 2nd tier: correct, 3rd tier: sure), as misconception (specific combinations of tiers as reported by Peşman and Eryılmaz (2010)), and as incorrect (e.g. all "not sure" responses to 3rd tier, any response does not fit into misconception category but incorrect, neglecting confidence level). Incorrect answers also indicate "lack of knowledge" about the electricity concept.

Kruskal-Wallis tests indicated that there were significant differences in PR students' correct response scores ( $\chi^2$ (4, 296) = 30.146, p < .001), incorrect response scores ($\chi^2$ (4, 296) = 50.570, p < .001), and misconception scores ($\chi^2$ (4, 296) = 49.368, p < .001) across different academic years. Post-hoc comparisons are conducted by using the Bonferroni correction for multiple tests, indicating the mean rank of students' correct responses, incorrect responses, and misconception scores in the SECDT. The pairwise comparisons for correct, incorrect, and misconception scores across academic year clusters are displayed in Figure 2.

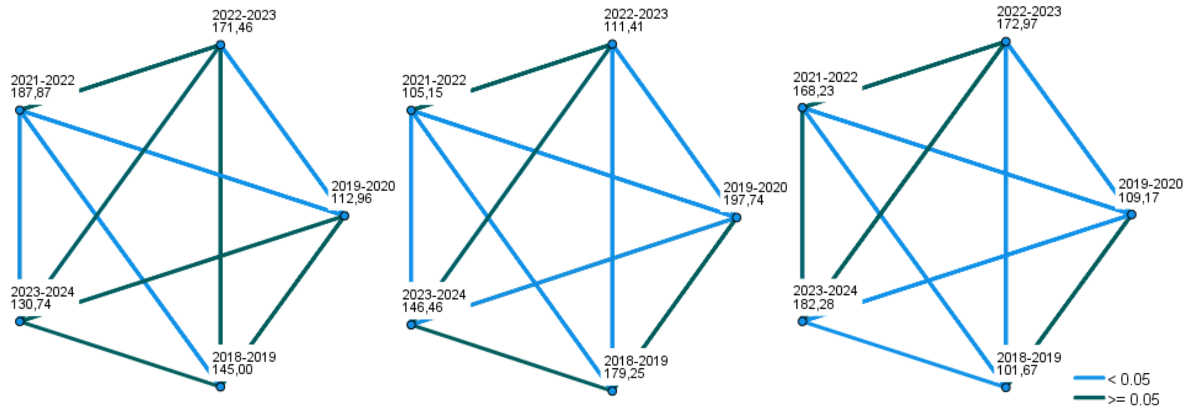



**Figure 2** Pairwise Comparisons Across Years of PR Students' (a) Correct Response Scores, (b) Incorrect Response Scores, and (c) Misconception Scores in the SECDT.

The SECDT correct response scores of PR students enrolled in the 2021-2022 academic year were significantly higher than those of students enrolled in the 2018-2019 cluster, p = 0.047, students enrolled in 2019-2020, p < .001, and students enrolled in 2023-2024, p < .001. On the other hand, the SECDT incorrect response scores of PR students enrolled in the 2021-2022 academic year were significantly lower than those of students enrolled in the 2018-2019 cluster, p < .001, students enrolled in 2019-2020, p < .001, and students enrolled in 2023-2024, p = 0.035. Even though PR students in the 2021-2022 cluster have higher correct and lower incorrect responses in the SECDT, their misconception scores were significantly higher than those of students enrolled in the 2018-2019 cluster, $p < .001$, and students enrolled in 2019-2020, $p < .001$.

The SECDT correct response scores of PR students enrolled in the 2022-2023 academic year were significantly higher than those enrolled in the 2019-2020 cluster, $p = 0.047$. Conversely, the SECDT incorrect response scores of PR students enrolled in the 2022-2023 academic year were significantly lower than those of students enrolled in the 2018-2019 cluster, $p = 0.001$, and students enrolled in the 2019-2020 cluster, $p < .001$. When the SECDT misconception scores of PR students enrolled in the 2022-2023 academic year were considered, they were significantly higher than those of students enrolled in the 2018-2019 cluster, $p < .001$, and in 2019-2020, $p < .001$.

## Discussions and Suggestions

The current study tracked health science students' conceptual understanding levels of simple electric circuits within five academic years to determine the trend of prevalent misconceptions brought to the university level. The results revealed a peak in 2021-2022 regarding better conceptual understanding level (see Figure 1 and Figure 2a). This might imply that these students were positively affected by the school closures during the COVID-19 pandemic. Attending online classes can create a sense of being in a physical classroom. When instructors use a video camera, students feel more motivated, and watching the instructor explain the topics helps them concentrate better (Al-Kumaim et al., 2021). Gökbulut (2021) identified a moderate positive association between the level of educational perception among higher education students and their readiness for mobile learning. This association was attributed to the fact that individuals can access information through mobile devices at any time, regardless of their location, thereby taking advantage of the benefits of distance education. The current study showed that this temporary effect due to a novelty (online education) has declined throughout the years. Moreover, the same effect was not observed for the 2022-2023 cluster affected by online education due to the Kahramanmaraş earthquakes. The difference may stem from the attendance not being mandatory.

However, considering misconception scores across the years (see Figure 2c), students of the 2018-2019 and 2019-2020 groups, who were not exposed to online education during high school, have fewer misconceptions than the other groups. Similarly, according to Koto and Ilhami (2023), students have been found to develop misconceptions about dynamic fluids following the shift to online learning during the COVID-19 pandemic. This can be attributed to reduced interaction between students and teachers, technical issues, and difficulty grasping key concepts.

The SECDT questions that were rarely answered correctly (Q2, Q3, Q5, Q9, and Q10) within five years indicate that students did not have a sound conceptual understanding related to how electric current flows through the series or parallel circuits, how the magnitude of electric current changes in a series of bulb connections, how the magnitude of electric current affected when changing the place of circuit elements, and lastly, how a bulb should be connected with a conducting wire and a cell to give light.

There is a thin line between having a good conceptual understanding and holding less misconception, and it cannot be stated that there is a linear relationship between them. The results of pairwise comparisons of PR students' performances over five years (see Figure 2) provide evidence for this. When the students' performances are evaluated from a misconception point of view, 7 out of 11 misconception models diagnosed by the SECDT were rarely detected within five years. These are the unipolar/sink model (M1), the attenuation model (M2), the shared current model (M3), the empirical model (M5), the parallel circuit misconception (M8), and the current flow as water flow misconception (M11). Even though students did not have a sound conceptual understanding of electric current, the aforementioned misconception models were not prevalent among students in different year clusters. We can conclude that students lack knowledge of these concepts.

In a prior analysis, Mackay and Hobden (2012) utilized diagrams to explore the preconceived notions of South African university students regarding electric circuits. The findings revealed that 34% of the "errors" stemmed from unipolar thinking. The researchers discussed that the unipolar model tends to resist prolonged instruction. However, in the current study, the percentage of PR students in each academic year cluster holding the unipolar/sink model was apparently lower than that proposed by Mackay and Hobden's (2012) findings. Tsai et al. (2007) suggested that the variation among studies might have resulted from various researchers' use of different diagnostic tools.

The study detected five prevalent misconception models between 2018 and 2023 clusters with percentages of more than 10%: the clashing current model (M4), the short circuit misconception (M6), the power source as constant current source model (M7), the sequential reasoning model (M9), and the local reasoning model (M10). Similar findings were reported in Manunure et al.'s study (2020), which identified the secondary students' misconceptions. Pretest results on the SECDT of control and experimental groups showed that the clashing current and short-circuit misconception models were the most prevalent misconceptions among students.

Moreover, these results are comparable to Peşman and Eryılmaz's (2010) and Aligo et al.'s (2021) findings. Peşman and Eryılmaz (2010) reported the most common misconception models among high school students as the shared current (M3), clashing current (M4), short circuit (M6), power supply as a constant current source (M7), and local reasoning (M10) models. In the current study, the shared current model was uncommon among students except for the 2022-2023 cluster (10.5%). As the trend of M3 declines through the following year, it is difficult to note that the shared current model is prevalent among PR students. In the latter study, the most common misconception model among high school students was found as the clashing current model (M4), which is followed by short circuit (M6), empirical rule (M5), shared current (M3), local reasoning (M10), and current flow as water flow (M11) models. In the same study, researchers examined the prevalence of misconceptions among science teachers of these high school students and found that the short circuit model (M6) was the most common misconception, followed by current flow as water flow (M11), clashing current (M4), shared current (M3), local reasoning (M10), and parallel circuit (M8) misconception models. This finding may be projected in the current study, which shows that potential source of common misconceptions might be the K-12 teachers' misconceptions. Further research may focus on longitudinal studies to update the sources of alternative conceptions starting from early years in new era.

Taslidere and Yıldırım (2023) found that many university students held onto the misconception of the clashing current model (M2) before and after receiving instructions. The researchers concluded that providing conceptual change texts enhanced with concept cartoons could improve primary preservice teachers' understanding of simple electricity. However, even after instruction, students were still inclined to think that the clash between positive and negative electricity in an electrical device made it operate. The current study's findings also showed that, over the years, the clashing current model has been quite persistent among students, implying that this might not have been addressed adequately through K-12 education. These results might also imply that there is not a single solution to remediate each misconception, and that new teaching strategies should be developed and implemented in the early stages of education, especially for the clashing current model.

In conclusion, PR students hold various misconceptions that might impede their practical work, which should be addressed in introductory physics courses. Five-year data shows minimal changes occurred, and the prevalent misconceptions did not change over time. These findings can help educators in designing their courses by explicitly focusing on these misconceptions to promote a better conceptual understanding of electricity.

**Compliance with Ethical Standards**

*Disclosure of potential conflicts of interest*

The author declares no competing interests.

*Funding*

The Başkent University Research Fund supported this study.

*CRediT author statement*

This study has a single author, and according to ethical rules, the responsible author carried out all research processes.

*Research involving Human Participants and/or Animals*

This retrospective research was conducted in accordance with ethical principles and rules. It was approved by the Başkent University Institutional Review Board and Ethics Committee (Project no: KA23/284).

## Fizyoterapi Lisans Öğrencilerinin Basit Elektrik Devrelerine İlişkin Kavramsal Anlama Düzeylerinin ve Kavram Yanılgılarının 2018-2023 Döneminde İncelenmesi

**Özet:**

Bu çalışmanın amacı, 2018-2023 yılları arasında, fizyoterapi lisans öğrencilerinin basit elektrik devreleri ile ilgili kavramsal anlama düzeyleri ile fizik derslerine getirdikleri kavram yanılgılarının ortaya çıkartılmasıdır. Bu retrospektif çalışmada açıklayıcı bir araştırma metodolojisi benimsenmiştir. Katılımcılar, amaca uygun örnekleme yöntemi kullanılarak, 2018-2023 yılları arasında bir vakıf üniversitesinde Fizik II dersine kayıt olan fizyoterapi ve rehabilitasyon lisans öğrencilerinden seçilmiştir. Yaşları 18 ila 22 arasında değişen toplam 296 öğrenci (209 kadın ve 87 erkek) tarafından gönüllü olarak doldurulan Basit Elektrik Devreleri Tanı Testi (BEDTT) sonuçları çalışmaya dahil edilmiştir. Öğrencilerin BEDTT'ye verdikleri yanıtlar MS Excel programı kullanılarak analiz edilmiş ve IBM PASW18 programı ile ileri analizler yapılmıştır. Her bir kavram yanılgısı kategorisinin frekansı, öğrencilerin o kategoriyle ilgili tüm sorulara verdikleri yanıtları toplayarak hesaplanmıştır. Farklı akademik yıllarda oluşturulan öğrenci gruplarının BEDTT puanları açısından istatistiksel olarak farklı olup olmadığını belirlemek için bağımsız örneklem Kruskal Wallis testi ve post hoc analizi (ikili karşılaştırmalar) yapılmıştır. Sonuçlar, öğrencilerin beş yıl boyunca beş ortak kavram yanılgısına sahip olduğunu göstermiştir: Bunlar, çarpışan akımlar modeli, kısa devre kavram yanılgısı, sabit akım kaynağı olan güç kaynağı modeli, sıralı muhakeme (akıl yürütme) ve bölgesel muhakeme modelleridir. Ayrıca, çalışma 2021-2022 yıllarında öğrencilerin diğer yıllara kıyasla daha iyi bir kavramsal anlama düzeyine sahip olduklarını ortaya koymuştur. Beş yıllık veri, genel olarak kavram yanılgılarının çok az değiştiğini, yaygın kavram yanılgılarının ise zaman içinde aynı kaldığını göstermiştir. Bu bulgular göz önünde bulundurularak, eğitimcilerin derslerini tasarlarken özellikle bu kavram yanılgılarına odaklanarak elektrik konusunun kavramsal olarak daha iyi anlaşılmasını sağlamalarına yardımcı olabileceği düşünülmektedir.

*Anahtar kelimeler*: Fizikçi olmayanlar için fizik, fizik dışı bilim dalları, sağlık bilimleri, kavramsal anlama düzeyi, kavram yanılgısı, fizyoterapi, basit elektrik devreleri.

## References


Al-Kumaim, N. H., Alhazmi, A. K., Mohammed, F., Gazem, N. A., Shabbir, M. S., & Fazea, Y. (2021). Exploring the impact of the COVID-19 pandemic on university students' learning life: An integrated conceptual motivational model for sustainable and healthy online learning. *Sustainability*, *13*(5), 2546. https://doi.org/10.3390/su13052546

Aligo, B. L., Branzuela, R. L., Faraon, C. A. G., Gardon, J. D., & Orleans, A. V. (2021). Teaching and learning electricity—A study on students' and science teachers' common misconceptions. *Manila Journal of Science*, *14*, 22-34. https://www.dlsu.edu.ph/wp-content/uploads/pdf/research/journals/mjs/MJS14-2021/issue-1/MJS14-3-2021-gardon-et-al.pdf

Altun, S. (2009). Determination of prospective science teachers' misconceptions of simple electric circuits issue with three- stage test. *Journal of Bayburt Education Faculty*, *4*(1), 72-79. https://dergipark.org.tr/en/pub/befdergi/issue/23159/247373

Arı, Ü., Peşman, H., & Baykara, O. (2017). Interaction of effect upon remediating prospective science teachers' misconceptions by guidance level in inquiry teaching with science process skills. *Bartın University Journal of Faculty of Education*, *6*(1), 304-321. https://doi.org/10.14686/buefad.263895

Ateş, S. & Polat, M. (2005). The effects of learning cycle method on removing misconceptions related to electric circuits. *Hacettepe University Journal of Education*, *28*(28), 39-57. https://dergipark.org.tr/en/pub/hunefd/issue/7808/102418

Caleon, I., & 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. https://doi.org/10.1080/09500690902890130

Caruana, E. J., Roman, M., Hernández-Sánchez, J., & Solli, P. (2015). Longitudinal studies. *Journal of Thoracic Disease*, *7*(11), E537-40. https://doi.org/10.3978/j.issn.2072-1439.2015.10.63

Chambers, S. K., & Andre, T. (1997). Gender, prior knowledge, interest, and experience in electricity and conceptual change text manipulations in learning about direct current. *Journal of Research in Science Teaching*, *34*(2), 107–123. https://doi.org/10.1002/(SICI)1098-2736(199702)34:2<107::AID-TEA2>3.0.CO;2-X

Cohen, R., Eylon, B., & Ganiel, U. (1983). Potential difference and current in simple electric circuits: A study of students' concepts. *American Journal of Physics*, *51*(5), 407-412. https://doi.org/10.1119/1.13226

Council of Higher Education [CoHE] (2023, March 30). *YÖK Başkanı Özvar, 2022-2023 Eğitim ve Öğretim Yılı Bahar Dönemi'ne ilişkin alınan yeni kararları açıkladı*. https://www.yok.gov.tr/Sayfalar/Haberler/2023/yok-baskani-ozvar-2022-2023-egitim-ogretim-bahar-donemi-ne-iliskin-alinan-yeni-kararlari-acikladi.aspx

Çepni, S. & Keleş, E. (2006). Turkish students' conceptions about the simple electric circuits. *International Journal of Science and Mathematics Education*, *4*, 269-291. https://doi.org/10.1007/s10763-005-9001-z

Duman, M. Ş., & Avcı, E. (2014). Fen ve teknoloji eğitiminde kavram yanılgıları üzerine 2003-2013 yılları arasında yapılmış çalışmaların değerlendirilmesi. *Fen Bilimleri Öğretimi Dergisi*, *2*(2), 67-82. https://dergipark.org.tr/tr/pub/fbod/issue/71981/1157958

Dupin, J. J., & Johsua, S. (1987). Conceptions of French pupils concerning electric circuits: Structure and evolution. *Journal of Research in Science Teaching*, *24*(9), 791-806. https://doi.org/10.1002/tea.3660240903

Engelhardt, P. V., & Beichner, R. J. (2004). Students' understanding of direct current resistive electrical circuits. *American Journal of Physics*, *72*(1), 98-115. https://doi.org/10.1119/1.1614813

Eryılmaz, A., Özdemir Ö. F., Yıldırım, U., Eraslan, F., Eryurt, K., Garip, B., & Demirtaş, D. (2016, September 28-30). *Lise öğrencilerinin basit elektrik devreleri ve geometrik optik konusundaki kavram yanılgıları* [Conference presentation abstract]. 12. Ulusal Fen Bilimleri ve Matematik Eğitimi Kongresi, Trabzon, Türkiye. https://open.metu.edu.tr/handle/11511/88149

Espejo-Antúnez, L., Fernández-Morales, C., Hernández-Sánchez, S., Cardero-Durán, M. d. l. Á., Toledo-Marhuenda, J. V., & Albornoz-Cabello, M. (2022). The impact on the stress-associated autonomic response of physiotherapy students receiving interferential current in an electrotherapy training session. *International Journal of Environmental Research and Public Health*, *19*(20), 13348. https://doi.org/10.3390/ijerph192013348

Fredette, N., & Lochhead, J. (1980). Student conceptions of simple circuits. *The Physics Teacher*, *18*(3), 194-198. https://doi.org/10.1119/1.2340470

Gökbulut, B. (2021). Distance education and mobile learning from the perspective of distance education students. *Educational Technology: Theory and Practice*, *11*(1), 160-177. https://doi.org/10.17943/etku.797164

Haladyna, T. M. (2004). *Developing and validating multiple-choice test items* (3rd ed.). Routledge. https://doi.org/10.4324/9780203825945

Hammer, D. (1996). More than misconceptions: Multiple perspectives on student knowledge and reasoning, and an appropriate role for education research. *American Journal of Physics*, *64*(10), 1316–1325. https://doi.org/10.1119/1.18376

Kaltakci-Gurel, D., Eryilmaz, 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. https://doi.org/10.12973/eurasia.2015.1369a

Kaltakci-Gurel, D., Eryilmaz, A., & McDermott, L. C. (2017). Development and application of a four-tier test to assess pre-service physics teachers' misconceptions about geometrical optics. *Research in Science & Technological Education*, *35*(2), 238-260. https://doi.org/10.1080/02635143.2017.1310094

Karakuyu, Y. & Tüysüz, C. (2011). Misconceptions in electricity and conceptual change strategy. *Gaziantep University Journal of Social Sciences*, *10*(2),867-890. https://dergipark.org.tr/en/pub/jss/issue/24242/257010

Kollöffel, B., & de Jong, T. (2013). Conceptual understanding of electrical circuits in secondary vocational engineering education: Combining traditional instruction with inquiry learning in a virtual lab. *Journal of Engineering Education*, *102*(3), 375-393. https://doi.org/10.1002/jee.20022

Koto, I., & Ilhami, D. (2023). High-school students' conceptual understanding of fluid dynamics following online learning during the coronavirus pandemics. In M. L. Firdaus & A. Defianti (Eds.), *Advances in social science, education, and humanities research: Vol. 718. Mathematics and Science Education International Seminar 2021* (pp. 117-125). Atlantis Press. https://doi.org/10.2991/978-2-38476-012-1_16

Mackay, J., & Hobden, P. (2012). Using circuit and wiring diagrams to identify students' preconceived ideas about basic electric circuits. *African Journal of Research in Mathematics, Science and Technology Education*, *16*(2), 131-144. https://hdl.handle.net/10520/EJC126045

Manunure, K., Delserieys, A., & Castéra, J. (2020). The effects of combining simulations and laboratory experiments on Zimbabwean students' conceptual understanding of electric circuits. *Research in Science & Technological Education*, *38*(3), 289-307. https://doi.org/10.1080/02635143.2019.1629407

Martellucci, J. (2015). Basic concepts in electricity and electrotherapy. In J. Martellucci (Ed.), *Electrical stimulation for pelvic floor disorders* (pp. 61-74). Springer. https://doi.org/10.1007/978-3-319-06947-0_4

Mbonyiryivuze, A., Yadav, L. L., & Amadalo, M. M. (2019). Students' conceptual understanding of electricity and magnetism and its implications: A review. *African Journal of Educational Studies in Mathematics and Sciences, 15*(2), 55-67. https://doi.org/10.4314/ajesms.v15i2.5

McDermott, L. C., & Shaffer, P. S. (1992). Research as a guide for curriculum development: An example from introductory electricity. Part I: Investigation of student understanding. *American Journal of Physics*, *60*(11), 994-1003. https://doi.org/10.1119/1.17003

Melzer, A., Cochran, S., Prentice, P., MacDonald, M. P., Wang, Z., & Cuschieri, A. (2012). The importance of physics to progress in medical treatment. *The Lancet*, *379*(9825), 1534-1543. https://doi.org/10.1016/S0140-6736(12)60428-0

Mulhall, P., McKittrick, B., & Gunstone, R. (2001). A perspective on the resolution of confusions in the teaching of electricity. *Research in Science Education, 31*, 575-587. https://doi.org/10.1023/A:1013154125379

Neidorf, T., Arora, A., Erberber, E., Tsokodayi, Y., & Mai, T. (2020). Review of research into misconceptions and misunderstandings in physics and mathematics. In *Student Misconceptions and Errors in Physics and Mathematics: Exploring data from TIMSS and TIMSS Advanced (IEA Research for Education)* (Vol. 9). Springer. https://doi.org/10.1007/978-3-030-30188-0_2

Ozdemir, E., Coramik, M., & Urek, H. (2020). Determination of conceptual understanding levels related to optics concepts: the case of opticianry. *International Journal of Education in Mathematics, Science and Technology*, *8*(1), 53-64. https://doi.org/10.46328/ijemst.v8i1.728

Önder, F., Şenyigit, Ç., & Sılay, İ. (2017). The effects of misconceptions on pre-service teachers' ability to constructing simple electric circuits. *European Journal of Physics Education*, *8*(1), 1-10. https://files.eric.ed.gov/fulltext/EJ1217022.pdf

Özmen, K. (2019, July 1-5). *Diagnosis of undergraduate physiotherapy students' misconceptions about simple electric circuits* [Conference presentation abstract]. The GIREP-ICPE-EPEC-MPTL 2019 Conference, Budapest, Hungary. https://edu.u-szeged.hu/ttkcs/sites/default/files/uploads/files/news/girep_icpe_epec_mptl_conference_abstract_book.pdf

Özmen, K. (2022, January 14-16). *Odyoloji bölümü öğrencilerinin elektrik devreleriyle ilgili kavram anlamlarının değerlendirilmesi* [Conference presentation abstract]. V. Ulusal Fizik Eğitimi Kongresi (UFEK), İstanbul, Türkiye. https://ufek2022.bogazici.edu.tr/V-UFEK-ozet-kitapcigi.pdf

Özmen, K. (2024). Health science students' conceptual understanding of electricity: Misconception or lack of knowledge?. *Research in Science Education*, *54*, 225–243. https://doi.org/10.1007/s11165-023-10136-3

Peşman, H. (2005). *Development of a three-tier to assess ninth grade students` misconceptions about simple electric circuits* (Publication No. 166885) [Master's thesis, Middle East Technical University]. Council of Higher Education Thesis Center.

Peşman, H., & Eryılmaz, A. (2010). Development of a three-tier test to assess misconceptions about simple electric circuits. *The Journal of Educational Research*, *103*(3), 208-222. https://doi.org/10.1080/00220670903383002

Quezada-Espinoza, M., Dominguez, A., & Zavala, G. (2023). How difficult are simple electrical circuit conceptions?. New findings. *European Journal of Educational Research*, *12*(3), 1269-1284. https://doi.org/10.12973/eu-jer.12.3.1269

Ramma, Y., Bholoa, A., Watts, M., & Nadal, P. S. (2017). Teaching and learning physics using technology: Making a case for the affective domain. *Education Inquiry*, *9*(2), 210–236. https://doi.org/10.1080/20004508.2017.1343606

Reiner, M., Slotta, J. D., Chi, M. T. H., & Resnick, L. B. (2000). Naive physics reasoning: A commitment to substance-based conceptions. *Cognition and Instruction*, *18*(1), 1-34. https://doi.org/10.1207/S1532690XCI1801_01

Rezigalla, A. A. (2020). Observational study designs: Synopsis for selecting an appropriate study design. *Cureus*, *12*(1), e6692. https://doi.org/10.7759/cureus.6692.

Shipstone, D. M. (1984). A study of children's understanding of electricity in simple DC circuits. *European Journal of Science Education*, *6*(2), 185-198. https://doi.org/10.1080/0140528840060208

Taslidere, E., & Yıldırım, B. (2023). Effect of conceptual change–oriented instruction on students' conceptual understanding and attitudes towards simple electricity. *International Journal of Science and Mathematics Education*, 21, 1567–1589. https://doi.org/10.1007/s10763-022-10319-w

Treagust, D. F. (1988). Development and use of diagnostic tests to evaluate students' misconceptions in science. *International Journal of Science Education*, *10*(2), 159-169. https://doi.org/10.1080/0950069880100204

Tsai, C. H., Chen, H. Y., Chou, C. Y., & Lain, K. D. (2007). Current as the key concept of Taiwanese students' understandings of electric circuits. *International Journal of Science Education*, *29*(4), 483-496. https://doi.org/10.1080/09500690601073327

Turgut, Ü., Gürbüz, F., & Turgut, G. (2011). An investigation 10th grade students' misconceptions about electric current. *Procedia-Social and Behavioral Sciences, 15,* 1965-1971. https://doi.org/10.1016/j.sbspro.2011.04.036

Uwamahoro, J., Ndihokubwayo, K., Ralph, M., & Ndayambaje, I. (2021). Physics students' conceptual understanding of geometric optics: Revisited analysis. *Journal of Science Education and Technology, 30,* 706–718. https://doi.org/10.1007/s10956-021-09913-4

Ünal, C. (2022). Pre-service science teachers' processes of establishing simple electric circuits. *Participatory Educational Research, 9*(1), 271-284. https://doi.org/10.17275/per.22.15.9.1

Vosniadou, S. (1994). Capturing and modeling the process of conceptual change. *Learning and Instruction, 4*(1), 45-69. https://doi.org/10.1016/0959-4752(94)90018-3

Yeltekin-Atar, B. Ş., Aykutlu, I., & Bayrak, C. (2021). An examination of studies on minconceptions in physics education conducted in the last 10 years in Turkey. *Journal of Kazım Karabekir Education Faculty, 42,* 304-323. https://doi.org/10.33418/ataunikkefd.831817

Yıldırım, H. İ., Yalçın, N., Şensoy, Ö. & Akçay, S. (2008). A survey study of 6th, 7th and 8th grade students' misconceptions in electric current. *Kastamonu Education Journal, 16*(1), 67-82. https://dergipark.org.tr/en/download/article-file/819072

Necatibey Faculty of Education Electronic Journal of Science and Mathematics Education
Vol. 18, Issue 2, December 2024, pp. 430-459. ISSN: 1307-6086

Research Article

# Classroom Teachers' Opinions on Game-Based Assessment in Mathematics Lesson: A Phenomenological Research*

**Ramazan DİVRİK[1], Hilal Nur TOPAL[2]**

[1] Trakya University, Faculty of Education, Türkiye, ramazandivrik@trakya.edu.tr, http://orcid.org/0000-0002-7126-7664

[2] Ministry of National Education, Cumhuriyet Primary School, Türkiye, hlnrglc@outlook.com, http://orcid.org/0009-0009-3441-5152

Received: 26.09.2024 Accepted: 13.11.2024

Doi: https://doi.org/10.17522/balikesirnef.1556395

---

*Abstract* - The study aims to gather the opinions of classroom teachers about the reflections of game-based assessment in mathematics lessons in primary school. The study followed a phenomenological design, which is one of the qualitative research methods. The participants consisted of eight classroom teachers working in public schools in Edirne and Kırklareli provinces in the 2023-2024 academic year. Participants were selected using convenience sampling, a type of purposive sampling method. The study data were collected using a semi-structured interview with six questions. The data were analyzed using content analysis. Based on the findings, game-based assessment was seen as a positive development. For teachers, it allowed long-term observation and monitoring of gains that turned into behaviors. For students, it fostered learning through fun, a desire to win, active participation, competition experience, and reduced exam stress. It also contributed to more enjoyable lessons and helped reduce negative perceptions of the subject. For teachers, challenges included the inability to make unbiased evaluations and low reliability. For students, negative aspects included lack of attention, focusing too much on fun, not following the rules of the game, and ignoring instructions. In terms of the mathematics lesson itself, issues included falling behind on the material and difficulty in creating games suitable for each learning objective. To address the challenges faced during game-based assessment in mathematics lessons, several strategies were suggested, including defining clear evaluation criteria, fostering cooperation and active participation, aligning objectives, following instructions, preparing a guideline, and ensuring material completion.

*Keywords*: Game-based assessment, primary school mathematics lesson, teacher views, phenomenology.

------------

Corresponding author: Ramazan DİVRİK, ramazandivrik@trakya.edu.tr

---

* This study was presented as an abstract at the 16th National Science and Mathematics Education Congress (UFBMEK2024).

## Introduction

Assessment and evaluation processes are critical to ensure mathematics teaching success and improve students' learning processes. These processes are vital in determining how much knowledge students acquire, which skills they develop, and in which areas they need support (Birgin & Gürbüz, 2008; Kutlu et al., 2010). Effective assessment and evaluation in education requires a systematic approach to accurately determine student performance and develop teaching strategies (Karasar, 2005). These processes generally include observation and evaluation stages. Comparisons made according to the criteria determined after observation using specific tools are necessary to evaluate the student's development and the effectiveness of the educational process (Turgut & Baykul, 2019).

Modernizing education and training processes has brought student-centered and innovative approaches. In this context, new approaches that observe the student's development at every stage and care about the process have been adopted by moving away from traditional evaluation methods (Acat & Uzunkol, 2010). These reforms in education systems have made assessment processes more dynamic and student-oriented (Pellegrino et al., 2001). This transformation includes increasing teachers' assessment competencies and improving students' academic achievement (Looney, 2009).

Assessment and evaluation processes guide teachers by determining students' knowledge and skill levels (Black & Wiliam, 1998). Therefore, appropriate methods that vary according to the course, subject, and outcome should be selected for assessment. In traditional assessments, students are evaluated with a single exam after a long period of education and teaching, disregarding their status. However, efforts have been made to change this, and the student's performance during the learning process has become significant. In contrast to the evaluation that emphasizes only paper and pencil when it comes to measurement and evaluation, new and various evaluations are made in which the student himself/herself, his/her interests, and abilities come to the forefront (Gelbal & Kelecioğlu, 2007). While one of these assessment methods, diagnostic assessment, shows in which areas students need help, formative assessment provides instant feedback during the learning process and enables updating teaching strategies (Shepard, 2000). Outcome-oriented assessment measures students' overall performance (Harlen, 2007). Studies conducted in Türkiye and internationally show that formative assessment increases student achievement (Black & Wiliam, 1998; Kutlu et al., 2010). International programs such as PISA and TIMSS use these assessments to compare education systems (Organisation for

Economic Cooperation and Development [OECD], 2019). In addition, Arnold (2011) emphasizes that formative assessment significantly contributes to student achievement when it is a feedback-based process.

In Türkiye, the Ministry of National Education (MoNE, 2023) has made various regulations to adapt to changes in assessment and evaluation processes. The regulation published in 2023 emphasized that students' academic and social development in grades 1, 2, 3, and 4 should be continuously monitored. For this purpose, it was requested that students be evaluated through game-based assessments and observation forms for participation in individual and group activities (MoNE, 2023). It also encouraged the adoption of process-oriented approaches in evaluating in-school work (MoNE, 2024a). In parallel with these developments, in the 2024 primary school mathematics curriculum, students are expected to be able to transform their skills into actions when necessary. For this purpose, supportive studies should be conducted to develop dispositions, which can be expressed as the predispositions needed to transform skills into actions. As mental patterns can be learned and developed, dispositions (curiosity, independence, empathy, assertiveness, playfulness, focus, original thinking, etc.) are triggering in exhibiting skills. The tendency of playfulness also comes to the fore as making learning a subject or behavior fun and enjoyable. In the curriculum, it is vital to create learning environments that will enable the emergence and support of dispositions and include learning experiences that offer the opportunity to observe dispositions concretely (MoNE, 2024a). Therefore, game-based assessments are necessary for providing concrete and enjoyable experiences to support the tendency of playfulness foreseen by the curriculum.

Game-based assessment is an essential innovation in education. While traditional assessment methods usually measure students' knowledge and skills with a single exam, game-based assessment can offer various advantages to students. Learning through games aims to increase the student's interest in the lesson, increase motivation, and make learning fun (Offenholley, 2012; Tayfur, 2019). Therefore, using games to assess mathematical skills may potentially provide realistic feedback to students, as it will contribute to a more fun and active assessment process.

Games support learning processes and provide opportunities to reduce students' anxiety about the lesson, relax, and have fun (Çil & Sefer, 2021; Offenholley, 2012; Tayfur, 2019). In these ways, games can effectively solve the perception of mathematics lessons as boring and intimidating. International literature emphasizes that game-based learning and assessment practices positively affect students' learning experiences (Gee, 2003; Papert, 1993). In particular, the potential of games to increase student motivation and make learning processes more effective is widely recognized. Game-based assessments in mathematics courses can reveal this potential, make students participate more in the lesson, and make learning processes more effective. Such assessment methods can help students experience mathematical concepts more concretely and reduce their prejudices against mathematics. Linking games to mathematics in this way, rather than seeing them as the opposite of work or leisure activity, will lead to greater student engagement and success (Offenholley, 2012).

Recently, it has been decided to abolish exams at the primary school level and replace them with game-based assessment in Türkiye (MoNE, 2023). This points to a new approach at the primary school level, unlike the preschool period. In this context, questions like "How can game-based assessment be applied in elementary school mathematics lessons?" "What form should such assessments take?" and "What are the difficulties encountered in practice?" constitute the rationale for this study. However, there have been many studies using play as an assessment tool at the international level (Chiu & Hsieh, 2017; Gomez et al., 2023; Kiili & Ketamo, 2018; Leong & Toh, 2021; Skillen et al., 2023) at the national level, it is seen that studies on game-based assessment have been conducted mainly at the preschool level (Çelik & Demirbaş, 2023; Işıkoğlu Erdoğan & Canbeldek, 2017; Okatan & Tagay, 2021). Therefore, this study, which was conducted to determine how game-based assessment can be an effective tool in mathematics teaching at the primary school level and how these practices can be improved, is considered necessary to provide essential findings for teachers and policymakers.

In the literature, it is known that many studies have been conducted on using games in education, and their effects have been examined using different variables. When these studies are examined, it is seen that games can be used as an assessment tool (Okatan & Tagay, 2021) and in value transfer (Gündüz et al., 2017), effective in the treatment of developmental and psychological disorders (Genç & Çakmak Tolan, 2021); and effective in variables such as problem-solving (Pintér, 2010; Şahin, 2019), concentration (Gözüm & Kandır, 2020), motivation (Yazıcıoğlu & Çavuş-Güngören, 2019), retention (Demir, 2016), attitude (Şanlıdağ & Aykaç, 2021), achievement (Dönmez et al., 2021; Ergül & Doğan, 2022). In addition, game-based activities have positive contributions such as fun, concretization, active participation, social interaction, time-saving, student motivation, permanent learning, attracting attention and instant feedback (Çil & Sefer, 2021; Demir, 2016; Gür & Kobak Demir, 2016; Katmada et al., 2014); it was evaluated that they would have positive effects on breaking the prejudice against mathematics lessons and increasing interest in the course (Demir, 2016; Özata & Coşkuntuncel, 2019). It has also been determined that the use of games in mathematics lessons has limitations such as not being suitable for crowded classes, not being applied due to lack of materials and environment in schools, taking too much time, and making classroom management difficult (Çil & Sefer, 2021; Ergül & Erşen, 2023; Gür & Kobak Demir, 2016; Özata & Coşkuntuncel, 2019).

In 2023, the Ministry of National Education amended the measurement and evaluation regulation: "In primary schools, the success of students is monitored by taking into account their developmental levels, their participation in teacher-guided lesson activities, observation forms for participation in individual and group activities, game-based assessments and measurement tools for fulfilling the assigned tasks. It is shown on the report card as "very good," "good," "sufficient," and "should be improved." (MoNE, 2023). Therefore, starting from the 2023-2024 academic year, the evaluation of students with game-based tools and the formation of report card opinions (very good, good, sufficient, should be improved) according to these results have brought up the issue of how the game will be used as an evaluation tool in mathematics lessons and what its effects will be. In this context, the purpose of this study is to obtain the opinions of classroom teachers about the reflections of game-based assessment in mathematics lessons in primary school, a recent change in the education system. The research questions were formulated as follows:

1. What are the opinions of classroom teachers about game-based assessment in mathematics lessons?

2. What are the opinions of classroom teachers about the difficulties encountered during game-based assessment in mathematics lessons?

3. What are the suggestions of classroom teachers for game-based assessment in mathematics lessons?

## Method

### Research Design

This study, in which teachers' views on game-based assessment in mathematics lessons were taken, was structured according to the phenomenological design, one of the qualitative research designs. Qualitative research involves a qualitative process in which qualitative data collection methods such as observation, interview, and document analysis are carried out and examine events and phenomena in their reality (Yıldırım & Şimşek, 2021, p. 37). In phenomenology, which is based on personal experiences, a phenomenon is identified, and people's experiences, perceptions, and the meanings they attribute to this phenomenon are tried to be understood (Baş & Akturan 2017, p. 87). In other words, phenomenology investigates how people experience an existing situation, what meaning they attribute to it, and how it appears in its nature (Patton, 2018, p. 104). This study considered that the teachers participating in the research had experience with play-based assessment. Interviews were conducted with teachers working in public schools in the second semester of the academic year to obtain their opinions on this new practice. In addition, by conducting interviews with teachers working at different grade levels, the experiences and opinions of these teachers on game-based assessment for all grade levels at the primary school level were analyzed, and it was tried to reveal how they attributed meaning to the determined phenomenon.

**Participants**

The study participants comprised eight classroom teachers working in public schools in Edirne and Kırklareli provinces in the 2023-2024 academic year. The fact that one of the study's authors is a classroom teacher played a significant role in identifying and selecting the participant teachers. As a result of the author's conducting a series of interviews in his school and at the point of identifying teachers with experience in this practice, the research was carried out by identifying teachers who had experience in game-based assessment and who volunteered to participate in the study. Again, care was taken to ensure that teachers worked at all grade levels in determining the teachers participating in the study. There was also one teacher with whom a pilot interview was conducted before teacher selection. As a result of the interview with this teacher, it was evaluated that it was essential to include 4th-grade teachers in the study, and three 4th-grade teachers were included in the study. The reason why 4th-grade teachers were specifically selected is that written exams will no longer be used to evaluate students in the interview's academic year. For these reasons, the convenience sampling method, one of the purposeful sampling methods, was used to determine the study participants. The convenience sampling method, preferred for speeding up the research and being practical, generally includes the immediate environment the researcher can easily reach. Therefore, it is a way that saves time for the researcher (Yıldırım & Şimşek, 2021, p. 121). The reason for choosing this method is that one of the authors of this study is a teacher. The study included teachers whom the researcher knew in the school where the researcher worked and in her immediate surroundings and who volunteered to participate in the study. Information about the participants is presented in Table 1.

**Table 1** Characteristics of Participants

| Teacher | Gender | Education level | Experience | Occupational title | Age | Class of service |
|---|---|---|---|---|---|---|
| T1 | Male | Master's degree | 7 | Teacher | 33 | 4th grade |
| T2 | Female | Master's degree | 10 | Expert teacher | 38 | 4th grade |
| T3 | Female | Master's degree | 8 | Expert teacher | 34 | 2nd grade |
| T4 | Female | Bachelor's degree | 7 | Teacher | 32 | 3rd grade |
| T5 | Female | Bachelor's degree | 16 | Expert teacher | 43 | 3rd grade |
| T6 | Female | Bachelor's degree | 10 | Expert teacher | 36 | 1st grade |
| T7 | Male | Bachelor's degree | 8 | Teacher | 34 | 4th grade |
| T8 | Female | Bachelor's degree | 8 | Teacher | 34 | 2nd grade |

Table 1 shows two male and six female teachers participating in the study. Three of the teachers have master's degrees, and five have bachelor's degrees. The teachers' professional experience ranged between 7 and 16 years; four of them were expert teachers, and four were teachers. Their ages ranged between 32 and 43. One of the teachers is a first-grade teacher, two are second-grade teachers, two are third-grade teachers, and three are fourth-grade teachers. In addition, the teachers who participated in the interview were coded as T1 and T2, and these codes were used in the quotations. For example, T1 represents the number one teacher interviewed.

## Data Collection

The study data were collected using a semi-structured interview with six questions. As in this study, if the reflection of an issue on people, their thoughts and feelings on that issue are tried to be determined, the most accurate answer will be given by those people. For this reason, a semi-structured interview form was used to obtain answers to open-ended questions in a natural and free environment (Türnüklü, 2000).

Both expert opinions were taken when preparing the interview form, and a pilot interview was conducted. For this purpose, a 10-question draft form covering the research problems was prepared first. This form was submitted to the opinion of an expert in mathematics education, and two questions that needed to be more suitable for the research were removed. Then, eight questions in this draft form were applied by conducting a pilot interview with a 4th-grade classroom teacher. During the interview, it was seen that two questions could be evaluated within the other six questions. In addition, it was decided that it would be appropriate for the teachers to be aware of the changes made in the Ministry of

National Education's measurement and evaluation regulation and to add personal information (gender, education level, experience, professional title, age, and grade) to the interview form. Sample questions in the interview form are: What are your thoughts about using game-based assessment in your mathematics lessons? What are the advantages of using game-based assessment in your mathematics lessons? What difficulties do you encounter while implementing game-based assessment activities in your mathematics lessons? What are your opinions on how game-based assessment can be done more qualitatively in your mathematics lessons?

After the interview form was prepared, appointments were made to conduct interviews with teachers who volunteered to participate in the study. Interviews were conducted by meeting at the interview place on the day and time of the appointment. Before starting the interview, the teachers were informed about the purpose of the research, that the answers they shared would not be shared with anyone, and that their names would be kept confidential. They were also informed that the interview would be recorded with a voice recorder, and the interviews began. The interviews lasted approximately 30 minutes. Short notes were taken during the interviews to serve as a reference for the researchers.

**Data Analysis**

In this study, the data were analyzed using the content analysis method, codes were extracted, and categories were created. When conducting content analysis, criteria should be determined depending on the research question, and the data should be compared internally or with an external criterion. The raw data obtained here should be systematically transformed into categories (Öğülmüş, 2019). Content analysis provides a closer approach to the numerical and more apparent dimension in quantitative research, one of the most obvious differences between quantitative and qualitative research. In other words, it helps to understand and explain the emerging themes and patterns when analyzing qualitative data. Qualitative content analysis consists of four steps: coding data, finding themes, organizing codes and themes, and describing and interpreting findings (Yıldırım & Şimşek, 2021, p. 240).

In the first step of this analysis, the second author initially identified 57 codes from the interview data, and together with the first author, they reduced this number of codes to 54. They also divided the codes into positive and negative expressions. In the second stage, categories were created by considering the sub-problems. Out of 54 codes, 26 codes determined for the first sub-problem were categorized in terms of teacher, student, and course by considering positive and negative expressions. The 14 codes determined for the second sub-problem were divided into subcategories, such as before, during, and after the implementation of this assessment, which is a new practice. For the third sub-problem, 14 codes were categorized into three subcategories: teachers, students, and course. The researchers placed the codes into the categories independently and then came together. The 7 codes identified for the first sub-problem, such as “lack of materials,” “lack of information,” and “lack of time,” were included in the other sub-problems. For the second sub-problem, it was decided to change the code “games do not include the learning outcomes” to the code “games cannot be prepared for each learning outcome.” It was decided that it would be appropriate to add the codes “application in the form of competition” and “combining learning outcomes” to the 14 codes determined for the third sub-problem. In addition, it was decided that it would be more appropriate to change the name of the subcategory of suggestions for the “lesson” in this category to “other.” In this way, the number of codes, initially determined as 57, decreased to 48 after the authors came together and made evaluations and the procedures in the analysis process were completed. After completing these procedures, the findings were presented with direct quotations from the teachers' opinions that were thought to explain the codes and categories in the best way. In the presentation of the findings of the codes and categories, the frequency values of the teachers' opinions for each code were included.

**Validity and Reliability**

The most important criterion for ensuring validity and reliability in research is to reach the same objective results and observe similar situations no matter how many people examine the research. In addition, the most crucial criterion in validity is that the measurement tool focuses on what it aims to measure and measures it in the best way. Reliability is a prerequisite for validity; however, it is not sufficient on its own (Öğülmüş, 2019). Unlike quantitative research on validity and reliability, qualitative research may be skeptical because it is not based on numerical values. However, more methods used in measurement increase validity and reliability. Among these methods, credibility (internal validity), transferability (external validity), consistency (internal reliability), and confirmability (external reliability) are methods to increase validity and reliability (Creswell, 2013).

In order to ensure credibility in the research, participant characteristics were presented in detail by making detailed descriptions from participant views. The researchers decided to transfer the participants' views into codes and categories. In order to ensure transferability, participants' views were analyzed individually and independently of each other and grouped under standard codes. Direct quotations from the participants' views that best explained each code were included. In order to ensure consistency, expert opinions and supervision were frequently obtained during the interviews with the teachers participating in the interviews. In addition, consistency was checked by constantly comparing the findings and interpretations with each other. Throughout the research, the expert suggestions were taken into consideration. In order to ensure confirmability, the researchers tried to reflect on how the teachers approached the phenomenon of game-based assessment with an objective approach. In order to control objectivity, the interview data obtained from the participants were recorded in writing.

## Results

### Classroom Teachers' Opinions on Game-Based Assessment

Regarding the first sub-problem of the study, the positive and negative opinions of classroom teachers on the use of game-based assessment in mathematics lessons were determined in three categories: teacher, student, and course.

The data on the positive and negative aspects of game-based assessment for teachers are presented in Table 2.

**Table 2** Positive and Negative Aspects of Game-Based Assessment for Teachers

| Category | Subcategory | Code | f |
|---|---|---|---|
| From the Teacher's Perspective | Positive aspects | Long-term observation | 5 |
| | | Seeing the acquisition that turns into behavior | 7 |
| | Negative aspects | Inability to make an objective assessment | 8 |
| | | Low reliability of the assessment | 8 |

When Table 2 is examined, in the subcategory of positive aspects, five teachers expressed the code of game-based assessment in mathematics lessons providing the opportunity to make long-term observations, and seven teachers expressed the code of providing the opportunity to see the acquisition that turns into behavior. Teacher opinions in this subcategory are exemplified below.

*T5: In the old system, when we did exams, we evaluated a long process in a short time; that is, with an exam, this was insufficient, in my opinion, but when we play* *games, we evaluate children in a process again, which allows us to observe them during the game.*

*T7: When they play games, I don't just say this to evaluate them; I sit and watch the children. I have an idea about their behavior. Thanks to the game, I had the opportunity to see how acquisition turns into behavior in children.*

In the negative aspects subcategory, the answers given by eight teachers formed the codes that made it difficult to make an objective evaluation and low the reliability of the evaluation. Some of the negative opinions are exemplified below.

*T1: Equal conditions and standard criteria are required for an unbiased evaluation, and it is tough to provide them. Because game assessment is open to all kinds of interpretations.*

*T2: Since we have yet to get an evaluation scale, everyone makes their own judgments.*

*This situation creates confusion and causes unfairness.*

Data on the positive and negative aspects of game-based assessment for students are presented in Table 3.

**Table 3** Positive and Negative Aspects of Game-Based Assessment for Students

| Category | Subcategory | Code | f |
|---|---|---|---|
| From the student's perspective | Positive aspects | Learning with fun | 8 |
| | | Willingness to win | 3 |
| | | Active participation | 7 |
| | | Gaining competition experience | 7 |
| | | Reducing exam stress | 5 |
| | Negative aspects | Lack of attention and interest | 3 |
| | | Focus on having fun | 7 |
| | | Failure to comply with game rules | 4 |
| | | Inability to follow instructions | 6 |
| | | Exclusion by peers | 5 |
| | | Inability to adapt to group games | 6 |

When Table 3 is analyzed, in the positive aspects subcategory, eight teachers expressed the code of learning by having fun, three teachers expressed the code of desire to win, seven teachers expressed the code of active participation, seven teachers expressed the code of gaining competition experience, and five teachers expressed the code of reducing exam stress. Teacher views on this subcategory are exemplified below.

*T2: Even the word "game" brings joy to children in the classroom. At least my class is like this. Children have fun and learn. Math lessons are fun.*

*T5: I already try to use competition in my class, especially with hardworking children. It makes them want to win even more.*

*T3: When playing games, even if the student is sick, he/she eventually becomes involved in the game, which is very nice.*

*T1: If a competitive environment can be established in the classroom, the lessons will always be more fun and faster in my teaching life. Children get excited when they feel the spirit of competition.*

*T8: I think stress and fear will decrease in my class if I even try to say, "No more exams; we will play games."*

In the negative aspects subcategory, the code for lack of attention and interest was formed by three teachers, the code for focusing on having fun was formed by seven teachers, the code for not following the game rules was formed by four teachers, the code for not following the instructions was formed by six teachers, the code for being excluded by peers was formed by five teachers, and six teachers formed the code for not adapting to group games. Some of the negative opinions are exemplified below.

*T2: Attention deficit affects the evaluation result if students act with the game's logic and focus only on having fun.*

*T7: Since the first thing that comes to children's minds when they think of games is to have fun, just like we do, there is a possibility that they focus on having fun and getting away from the learning environment.*

*T1: Just like the evaluation criteria are not common, a healthy evaluation can only be made when the rules of the games are followed. If a student gets a low score because of such a game, it may cause injustice.*

*T2: Students may have difficulty following instructions due to a lack of interest.*

*T3: In case of any exclusion by the group, the child may not be able to reflect this even in a subject he/she knows and may get a low grade. In other words, the group will have a negative effect on the evaluation of this child.*

*T4: There will necessarily be games to be played with the group rather than individually, and the group is an entirely different issue. Not every child can adapt.*

Table 4 presents the data on the positive and negative aspects of game-based assessment regarding mathematics lessons.

**Table 4** Positive and Negative Aspects of Game-Based Assessment for Mathematics Lesson

| Category | Subcategory | Code | f |
|---|---|---|---|
| In terms of mathematics lesson | Positive aspects | Conducting fun lessons | 8 |
| | | Breaking the prejudice against the course | 6 |
| | Negative aspects | Failure to keep up with the subjects | 8 |
| | | Failure to prepare appropriate games for each achievement | 8 |

When Table 4 is analyzed, in the subcategory of positive aspects, eight teachers expressed the code for the lessons being fun, and six teachers expressed the code for breaking the prejudice against the course. Teacher views on this subcategory subcategory are exemplified below.

*T4: Exam anxiety is a reality, and it affects even us. The effect of games on children is indisputable. Using the game at every stage means more fun lessons in the simplest form.*

*T7: Whenever we play games in math class, they do not want the lesson to end. In this way, even they want to do math. I think this shows that the game reduces prejudice.*

In the negative aspects subcategory, the answers given by eight teachers formed the codes of not keeping up with the subjects and not preparing games suitable for each outcome. Some of the negative opinions are exemplified below.

*T1: Since the game will take more time than normal written assessments and time must be allocated for each student individually when evaluating each student with the game, I think I will have problems keeping up with the subjects.*

*T3: Some of the objectives will be ignored because no matter how many games there are, I don't think there can be a game that will fit every objective in every lesson at every grade level.*

## Difficulties Encountered While Conducting Game-Based Assessment

Regarding the second subproblem of the study, the difficulties encountered by classroom teachers while conducting game-based assessments in mathematics lessons were grouped under three subcategories: before, during, and after the implementation.

**Table 5** Difficulties Encountered During Game-Based Assessment in Mathematics Lessons

| Category | Subcategory | Code | f |
|---|---|---|---|
| Difficulties encountered | Before the application | Lack of knowledge | 7 |
| | | Lack of an application guide | 8 |
| | | Failure to prepare appropriate games for each achievement | 6 |
| | | Lack of knowledge of evaluation criteria | 8 |
| | During the application | Lack of time | 4 |
| | | Lack of materials | 4 |
| | | The game does not match the level of the student | 3 |
| | | Fast lesson processing | 5 |
| | | Textbooks do not support game-based assessment | 7 |
| | After the application | Failure to evaluate personal success | 5 |
| | | Gains that cannot be assessed | 7 |
| | | Differences in scoring | 4 |
| | | Assessment not aligned with achievement | 4 |

When Table 5 is examined, among the codes before the application subcategory, the code for lack of information was formed with the opinions of seven teachers, the code for lack of an application guide was formed with the opinions of eight teachers, the code for not being able to prepare a game suitable for each outcome was formed based on the opinions of six teachers, and the code for not knowing the evaluation criteria was formed with the opinions of eight teachers. Some of the related teacher opinions are as follows.

*T4: Honestly, I learned about this innovation thanks to this interview. I had not received any information about implementing it at school or in any other way before.*

*T5: We are expected to produce games that include all the outcomes, but if each teacher tries to produce and implement easy and difficult games, there will be confusion and a waste of time. In other words, the biggest need right now is a rich guide full of games that include the learning outcomes for all of us.*

*T6: Evaluating each outcome or topic with games requires a long time. However, one day was enough for exams. I think it has become very difficult to complete the subjects or to complete the assessment in its entirety.*

*T8: Games are something that both students and teachers facilitate, but if we are going to evaluate and grade a student according to games, there should be specific criteria. Otherwise, it is a vast spectrum. This issue is inextricable for me.*

Among the codes in the subcategory during the application, the code for the shortage of time and materials was formed with the opinions of four teachers, the code for the fact that not every game appeals to the level of every student was formed with the opinions of three teachers, the code for fast lesson processing in order to apply the games was formed with the opinions of five teachers and the code for the fact that textbooks do not support game-based assessment was formed with the opinions of seven teachers. Some of the related teacher opinions are as follows.

*T1: In the past, allocating one class hour for exams while grading was enough, but now it takes a lot of time even to play a game. We had a lot of trouble with this.*

*T5: Games set up in a more material-rich environment will appeal to more senses, but the evaluation will not be qualified without material.*

*T4: Every game may only work for some students. While hardworking students may get bored in a game with straightforward content, we may not be able to evaluate low-level students in challenging games.*

*T8: I tried to teach the lesson faster by arguing that games should be designed and applied to include each outcome, but I cannot say that it was very productive.*

*T1: I find the number of games in the textbooks insufficient. Games are at the center of primary school children's lives. Therefore, more games should be included in the books.*

Among the codes in the “after the application” subcategory, the code for the inability to evaluate personal achievement was formed with the opinions of five teachers, the code for the gains that could not be evaluated was formed with the opinions of seven teachers, and the codes for the differences in scoring and the incompatibility of the evaluation with the achievement were formed with the opinions of four teachers. Some of the related teacher opinions are as follows.

*T1: Evaluating students during a game becomes difficult. We also set up games that emphasize a single person's performance, but inevitably, there are also group games, and here, I had to give points to the group.*

*T5: I had the most difficulty designing games to evaluate all the outcomes. There were some outcomes that I could not fully evaluate. There was inevitably a deficiency.*

*T2: I think I prepared the games with a little difficulty. But this was not so challenging for every student; some got high scores from easy games. We need common evaluation criteria and games for all of us. Otherwise, there will be scoring differences between teachers.*

*T8: I have some children who are very active and successful in class but cannot show themselves in games. I have seen this in socially weak children. They deserve a much better score than the one they got in the game.*

## Suggestions for Game-Based Assessment

Regarding the third subproblem of the study, classroom teachers' suggestions for using game-based assessment in mathematics lessons were grouped into three subcategories: suggestions for teachers, suggestions for students, and other suggestions.

**Table 6** Suggestions for Game-Based Assessment in Mathematics Lessons

| Category | Subcategory | Code | f |
|---|---|---|---|
| Suggestions | To teachers | Cooperation | 5 |
| | | Identify common evaluation criteria | 8 |
| | | Material completion | 3 |
| | | Identifying level-appropriate games | 2 |
| | | Group evaluation | 2 |
| | | Application in the form of a competition | 3 |
| | | Gain consolidation | 6 |
| | | Time planning | 4 |
| | To students | Separation into groups | 3 |
| | | Active participation | 7 |
| | | Following instructions | 5 |
| | | Cooperation | 5 |
| | Other | Application guide | 8 |
| | | Providing instructions | 8 |
| | | Material support | 6 |
| | | Supporting textbooks | 7 |

When Table 6 is analyzed, eight codes for teacher suggestions and four for student and other suggestions are found.

For teachers, the codes are as follows: Cooperation five, determining common evaluation criteria eight, completing materials three, identifying level-appropriate games two, group evaluation two, competition three, combining learning outcomes six, and time planning four. The opinions of the teachers in this category are exemplified below.

*T7: To tolerate the lack of a guide, it would be easier if the teachers at least determine games among themselves. Cooperation can be done through social platforms or among themselves to exchange games and evaluation criteria.*

*T2: Teachers can form groups and evaluate with standard criteria.*

*T3: The materials of the identified games can be listed, and the missing materials can be obtained in advance.*

*T3: Games that are neither difficult nor easy for the class level should be produced by adhering to the principle of student relativity.*

*T2: It is impossible to evaluate a student alone in games played with a group; students can be evaluated by making a general scoring.*

*T1: Games can be applied as competitions, and time can be saved.*

*T5: Since it will be challenging to find games for each outcome for assessment with games, close outcomes can be combined and measured in a single game.*

*T6: Assessment with games creates time problems, so a plan should be made and proceed accordingly.*

The codes for students are as follows: separation into groups, three; active participation, seven; following the instructions, five; and cooperation, five. The opinions of the teachers in this category are exemplified below.

*T4: Students will find it easier to apply the games by separating into groups. They will progress in cooperation.*

*T6: Producing games that appeal to the whole class should ensure active participation. Students' participation in the evaluation will ensure that the evaluation is qualified. Students should not say, "I don't want to play this game."*

*T8: To provide a practical and accurate assessment, we must ensure that children understand and apply the instructions correctly. Their full compliance with the instructions will ensure that the assessment is correct for them.*

*T5: Students who are more familiar with the games and lessons can help others and play games together.*

Other codes are as follows: support for the application guide eight, providing instructions seven, material support six, and supporting textbooks seven. The opinions of the teachers in this category are exemplified below.

*T1: Teachers can be guided by sending a guidebook or a list of games in another way.*

*T7: I would expect instructions to manage the process of implementing the games well. It would be less tiring. We could better plan the application time.*

*T3: The materials should be complete, and the class size should be appropriate in order to apply the designed games quickly and effectively.*

*T7: I need help finding textbooks for the game. If we make the game indispensable for so many lessons, it should be included more in the textbooks.*

## Conclusions and Discussion

Based on the study's findings, it was revealed that there are positive and negative aspects of game-based assessment in elementary school mathematics lessons in terms of teachers, students, and the lessons themselves.

According to the results obtained regarding the positive aspects of play-based assessment, game-based assessment was a positive development because it allowed teachers to make long-term observations and observe behavior during the game. This result is important in allowing teachers to make objective and permanent evaluations. The fact that teachers use games to get to know the child in the preschool period overlaps with our research results and indicates that games can be an evaluation tool (Çelik & Demirbaş, 2023; Okatan & Tagay, 2021; Skillen et al., 2023). Gür and Kobak Demir (2016) found that using games in mathematics lessons provides instant feedback and supports permanent learning, consistent with our research results. According to this result, it is understood that classroom teachers can use game-based assessments to closely observe their students' mathematics performances in mathematics lessons.

While game-based assessment is supported because it will contribute to students' social aspects, such as having fun during a game, desire to win, active participation, and gaining competition experience, it is seen as an application allowing students to escape exam stress. A study conducted with higher education students revealed that their exam anxiety decreased, and their exam performance increased significantly in game-based assessment (Mavridis & Tsiatsos, 2017). Using game-based activities is seen as having positive effects in providing active participation, learning with fun, and increasing social interaction between students (Çil & Sefer, 2021; Şentürk, 2020). Unlike traditional methods, measurement and evaluation have become a process today. This study investigated whether games can be used as a process assessment tool in mathematics lessons. Unlike traditional assessment and evaluation practices, process-based assessments give permanent and successful results (Ergül & Doğan, 2022). Therefore, teachers can use game-based assessments to distract students from exam stress and socialize them.

When examined in terms of mathematics lessons, it was evaluated that they would positively contribute to making boring and fearful mathematics lessons fun and breaking prejudice against the lesson. It has been revealed in related studies that using games in mathematics lessons develops a positive attitude in children, increases interest (Demir, 2016), and increases student motivation (Tayfur, 2019). Mathematics teachers stated that game-based applications would reduce prejudices against mathematics and thus make the mathematics lesson more concrete and fun by getting rid of abstractness (Özata & Coşkuntuncel, 2019) in the study in which the systems in our body unit of the science course was evaluated using game-based assessment, an increase in students' motivation and academic achievement was observed, indicating that game-based assessments can be used in other courses (Tayfur, 2019). Games are used in mathematics courses to develop mathematical skills beyond procedural skills, including mathematical reasoning (Olson, 2007), conceptual understanding (Clarke & Roche, 2010), and problem-solving (Karayol & Temel, 2018; Pintér, 2010; Russo et al., 2021). Based on the results obtained in this study, it is understood that games can be used in the assessment and evaluation processes of mathematics courses in primary school.

According to the results obtained regarding the negative aspects of game-based assessment in the study, it was stated that it would not be possible to conduct game-based assessments in mathematics lessons due to the lack of common criteria and application guidelines. It was revealed that this situation would lead to results that could not be evaluated objectively and reliably. This result shows that scoring by using teachers' personal opinions or by observing students' instant performances will vary from person to person. However, it is essential to evaluate each student's mathematics performance in a way that does not change according to objective criteria and personal opinion. Students need to see and complete their deficiencies according to their performance in terms of measurement and evaluation principles. In order to comply with these principles, the primary school mathematics curriculum requires teachers to use tools such as checklists, follow-up tests, observation forms, portfolios, performance tasks, and rubrics to assess learning outcomes (MoNE, 2024a). However, although teachers express positive opinions about alternative assessment and evaluation approaches, it is known that they do not use these tools and prefer written exams and tests instead (Karakuş, 2010). Since formative assessment positively affects mathematics achievement, attitudes toward mathematics, and recall, teachers should be supported using these tools (Tekin, 2010). Therefore, teachers should be encouraged to use assessment tools they cannot use personal judgment for; teachers should be supported with implementation guides to objectively carry out assessment and evaluation processes.

The student's inability to focus on the mathematics subject by focusing only on having fun emerged as a negative opinion. In addition, not following the instructions during a game, not adapting to group games, and not following the game's rules may prevent the evaluation process from being carried out in a qualified manner. It may cause students to be excluded by their peers. Even each student's performance cannot be fully observed because they cannot adapt to group games. Therefore, it would be helpful to use interview forms, checklists, rating scales, and rubrics to evaluate students' performances according to the criteria specified in the forms and provide feedback (MoNE, 2024a). Monitoring practices aimed at identifying and raising such learning deficiencies will give the student the idea of recovering, participating in the lesson, paying attention, and thus reminding them of their responsibility for learning (Hotaman, 2020). In addition, Zengin et al. (2017) determined that Kahoot! and Plickers software have positive contributions, such as being applicable in the formative assessment process of mathematics teaching, facilitating assessment, and saving time by providing detailed and instant data analysis, which shows that student's attention can be attracted with digital tools. Therefore, students for whom having fun is of secondary importance should be assessed with appropriate assessment tools that reveal their actual performance. Ketamo and Devlin's (2014) finding that students who play too many games experience major conceptual misunderstandings during the game also supports our research findings. In addition, related research also stated that real-time game-based assessment is not yet ready to replace PISA to assess students “at the global level.”

The lack of time for the subjects and the inability to prepare appropriate games for each subject were negative assessment features with games in mathematics lessons. Trying to assess each subject with a game, taking more time than planned, or not having a suitable game for each outcome will negatively affect the quality of the lesson. Skillen et al. (2023)

examined the effects of a game-based assessment tool developed to measure the mathematical competence of preschool children. While the developed game was successful in at least 80% of low-achieving children, a negative feature emerged due to the time-consuming implementation of the game. Similarly, the fact that the preparations for the activities and in-class applications take too much time and that the activities make classroom management difficult are defined as the adverse effects of game-based activities (Çil & Sefer, 2021). Preparing games in which more than one subject or outcome will be assessed with a single game and assessing students at appropriate times will minimize the negative aspects of game-based assessment.

The difficulties encountered during game-based assessment in primary school mathematics lessons were grouped into three categories: before, during, and after the implementation. According to the results obtained from these categories, it was concluded that the lack of knowledge of teachers before starting the application, the lack of an application guide, the inability to prepare a game suitable for each outcome, and the lack of predetermination of evaluation criteria by teachers led to the results that the evaluation results could not provide reliable results. Regarding using educational mathematics games in mathematics teaching, mathematics teachers stated that it was challenging to prepare games for each outcome (Özata & Coşkuntuncel, 2019). Similarly, pre-service classroom teachers failed to design games at a satisfactory level in the dimensions of "game rules," "goals and objectives," "outcomes and feedback," and "interaction" (Pilten et al., 2017). The fact that teachers know the games they will use before starting game-based assessment and are prepared for the materials, duration, scoring instructions, etc., required by these games will ensure that the assessment is carried out more healthily.

During the application, it was determined that the game took too much time, there were no materials suitable for the assessment, the game was not suitable for the level of the students, and the textbooks did not support game-based assessment, which would lead to the lessons to be taught quickly or to be completed before they were fully completed, so the assessment could not achieve its purpose and valid results could not be obtained. The limitations encountered regarding the lack of materials and time in game-based mathematics lessons can be considered similar results (Gür & Kobak Demir, 2016). Another study concluded that games in game-based mathematics activities took too much time and made classroom management difficult (Çil & Sefer, 2021). In addition, teachers stated that educational mathematics games make classroom management difficult; games take too much time and are difficult to use in crowded classes (Özata & Coşkuntuncel, 2019). Therefore, it can be considered an important variable that the necessary materials and textbook support should be completed in order for the game-based assessment to be carried out in a qualified manner and that the process should be carried out without being rushed during the evaluation.

After the implementation, teachers evaluated assessment with games negatively because personal assessment cannot be made when group games are played, there will be incomplete scoring because there is not a game suitable for every outcome, there will be differences in scoring according to the student or teacher, and scoring cannot be done in accordance with personal success due to the performance brought by group games. While it has been determined that classroom teachers mainly apply performance assessment, portfolio, project, and self-assessment techniques and that feedback is the most common performance indicator of the techniques applied, it is known that classroom teachers are inadequate in alternative measurement and evaluation (Özenç & Çakır, 2015). The results of this study show that having appropriate assessment tools and rubrics for group or individual games ready in advance will enable objective scoring after the implementation.

The suggestions of the interviewed teachers for game-based assessment in primary school mathematics lessons were grouped into three categories: teachers, students, and others. According to the results obtained from these categories, it was suggested that teachers cooperate in determining common games, sharing materials, determining common assessment criteria, and combining learning outcomes to make a qualified assessment with games in mathematics lessons. In addition, it is recommended that games appropriate for each student level be determined in the classroom, and games should be designed in the form of competitions or in a way that supports group assessment to save time. Çil and Sefer (2021) determined that classroom teachers need inexpensive and easy-to-prepare materials to implement game-based activities, which coincides with our research findings. In addition, allowing teachers to organize the outcomes and the time allocated to the outcomes more freely will reduce classroom teachers' concerns about time and use game-based mathematics activities more frequently (Çil & Sefer, 2021). Based on this result, it would be appropriate for teachers to determine appropriate games to evaluate all their students, cooperate to complete the necessary materials, and apply these games in groups or in the form of competitions.

The teachers' suggestions for students to be divided into groups, to actively participate in this process by acting according to the instructions, and to cooperate were presented by the teachers as suggestions for the students to assess with games in a more qualified way. This result shows that game-based assessment in mathematics lessons should be handled not only with the teacher dimension but also with the student dimension. Similarly, Skillen et al. (2023) supported our research results by stating that there is a need to take the opinions of experienced teachers and children about game-based assessment and to conduct studies that reveal the advantages or disadvantages of game-based assessment compared to traditional tests. Students exposed to assessment should take this process seriously and actively participate to demonstrate their performance accurately.

Apart from these suggestions, having a guide that shows the game list, materials, duration, application instructions, and in which achievements the game can be used; organizing the activity and assessment sections (such as theme assessment) in the textbooks in a way to support the games are presented as important suggestions for a more qualified assessment with games. The inclusion of game-based activities in the first-grade mathematics textbooks prepared in accordance with the 2024 primary school mathematics curriculum can be considered a positive development in this respect (MoNE, 2024b). However, it is among the responsibilities of teachers to determine the criteria for these activities to evaluate students' mathematical skills. The research results suggest that the Ministry of National Education should prepare a booklet for classroom teachers to use in mathematics lessons for each grade level. This booklet provides teachers with a detailed description of each stage, from the game they can use to the scoring key.

## Limitations

This study was conducted using a phenomenology design, one of the qualitative research designs. The research is limited to eight classroom teachers working in Edirne and Kırklareli provinces in the 2023-2024 academic year who volunteered to participate in the interview. The data obtained are limited to the semi-structured interview form containing six questions created for this study. The data obtained from interviews with each teacher is limited to approximately 30 minutes.

**Compliance with Ethical Standards**

*Disclosure of potential conflicts of interest*

The authors declare that they have no competing interests.

*Funding*

This research did not receive external funding.

*CRediT author statement*

The first author developed the initial idea and then designed the study. The first author managed the research process and guided the data collection. The second author collected all the data under the supervision of the first author. The first and second author worked together in all data analyses and writing the manuscript.

*Research involving Human Participants and/or Animals*

This study was carried out taking into account ethical rules. The participants were informed about the study and asked to sign a consent form voluntarily. The research was approved by the decision numbered 05/26 of the Trakya University Social and Human Sciences Research Ethics Committee at its meeting dated 29.05.2024. The data of the study was collected in the second semester of the 2023-2024 academic year.

## Matematik Dersinde Oyun Temelli Değerlendirmeye İlişkin Sınıf Öğretmenlerinin Görüşleri: Fenomenolojik Bir Araştırma

**Özet:**

Araştırmanın amacı ilkokulda oyun temelli değerlendirme yapmanın matematik dersindeki yansımalarına ilişkin sınıf öğretmenlerinin görüşlerini almaktır. Çalışma nitel araştırma desenlerinden olgu bilim (fenomenoloji) desenine göre yapılandırılmıştır. Çalışmanın katılımcıları 2023-2024 eğitim öğretim yılında Edirne ve Kırklareli illerindeki devlet okullarında görev yapan sekiz sınıf öğretmeninden oluşmaktadır. Katılımcılar amaçlı örnekleme yöntemlerinden kolay ulaşılabilir örnekleme yöntemiyle belirlenmiştir. Çalışmanın verileri altı sorudan oluşan yarı-yapılandırılmış görüşme formu ile toplanmıştır. Veriler içerik analizi yöntemiyle analiz edilmiştir. Elde edilen bulgular ışığında; öğretmenler açısından uzun süreli gözlem yapma ve davranışa dönüşen kazanımı görme; öğrenciler açısından eğlenerek öğrenme, kazanma isteği, aktif katılım, yarışma deneyimi kazanma ve sınav stresini azaltma; ders açısından derslerin eğlenceli yürütülmesi ve derse olan ön yargının kırılması kodları ile oyun temelli değerlendirme olumlu bir gelişme olarak değerlendirilmiştir. Tarafsız değerlendirme yapamama ve değerlendirme güvenirliğinin düşük olması öğretmenler açısından; dikkat ve ilgi eksikliği, eğlenmeye odaklanma, oyun kurallarına uymama ve yönergeleri uygulayamama öğrenciler açısından; konuların yetişmemesi ve her kazanıma uygun oyunun hazırlanamaması matematik dersi açısından olumsuz yönler olarak görülmüştür. Matematik derslerinde oyun temelli değerlendirme yapılırken karşılaşılan güçlüklerin giderilmesine yönelik; değerlendirme kriterlerinin belirlenmesi, işbirliğini ve aktif katılımı teşvik etme, kazanım birleştirme, yönergelere uyma, uygulama kılavuzu hazırlama ve materyal tamamlama gibi önerilerde bulunulmuştur.

*Anahtar kelimeler*: Oyun temelli değerlendirme, ilkokul matematik dersi, öğretmen görüşleri, fenomenoloji.

## References

Acat, M., & Uzunkol, E. D. (2010). Primary school teachers' opinions about assessment process of primary education curriculums. *Electronic Journal of Social Sciences, 9*(31), 1-27. https://dergipark.org.tr/tr/pub/esosder/issue/6145/82482

Arnold, I. (2011). John Hattie: Visible learning: A synthesis of over 800 meta-analyses relating to achievement. *International Review Education, 57*, 219-221. https://doi.org/10.1007/s11159-011-9198-8

Baş, T., & Akturan, U. (2017). *Sosyal bilimlerde bilgisayar destekli nitel araştırma yöntemleri.* (3rd ed.). Seçkin.

Birgin, O., & Gürbüz, R. (2008). Sınıf öğretmeni adaylarının ölçme ve değerlendirme konusundaki bilgi düzeylerinin incelenmesi. *The Journal of Selcuk University Social Sciences Institute,* (20), 163-179. https://dergipark.org.tr/en/pub/susbed/issue/61796/924236

Black, P., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education: Principles, Policy & Practice, 5*(1), 7-74. https://doi.org/10.1080/0969595980050102

Chiu, F. Y., & Hsieh, M. L. (2017). Role-playing game based assessment to fractional concept in second grade mathematics. *Eurasia Journal of Mathematics, Science and Technology Education, 13*(4), 1075-1083. https://doi.org/10.12973/eurasia.2017.00659a

Clarke, D., & Roche, A. (2010). The power of a single game to address a range of important ideas in fraction learning. *Australian Primary Mathematics Classroom, 15*(3), 18-23. https://files.eric.ed.gov/fulltext/EJ898705.pdf

Creswell, J. (2013). *Araştırma deseni: Nitel, nicel ve karma yöntem yaklaşımları* [Research design: Qualitative, quantitative and mixed methods approaches] (S. B. Demir, Trans. Ed.; 4th ed.). Eğiten Kitap.

Çelik, Ö., & Demirbaş, A. (2023). Examining the recognition and evaluation processes of preschool children according to teacher opinions: A qualitative research. *International Journal of Primary Education Studies, 4(*3), 170-183. https://doi.org/10.59062/ijpes.1400279

Çil, O., & Sefer, F. (2021). The investigation of elementary school teachers' opinions on game-based mathematics activities. *Trakya Journal of Education, 11*(3), 1366-1385. https://doi.org/10.24315/tred.814024

Demir, M. R. (2016). *Effect of teaching mathematics through different play types on 1. grade students' achievements and retention of learning levels* (Publication No. 436706) [Doctoral dissertation, Necmettin Erbakan University]. Council of Higher Education Thesis Center.

Dönmez, B., Dönmez, K. H., Kolukısa, Ş., & Yılmaz, S. (2021). The effect of using the teaching by game method in primary school mathematics course on attitudes and success. *Bozok International Journal of Sport Sciences*, *2*(2), 58-70. https://dergipark.org.tr/tr/pub/ubosbid/issue/85237/1508426

Ergül, E., & Doğan, M. (2022). The effect of game-based teaching mathematics on student achievement in elementary school. *Milli Eğitim Dergisi, 51*(235), 1935-1960. https://doi.org/10.37669/milliegitim.887654

Ergül, E., & Erşen, Z. B. (2023). Should primary school mathematics education be gamification or not? (the opinions of primary teachers). *The Journal of Turkish Educational Sciences*, *21*(1), 49-77. https://doi.org/10.37217/tebd.1173722

Gee, J. P. (2003). What video games have to teach us about learning and literacy. *Computer in Entertainment*, *1*(1), 1-4. https://doi.org/10.1145/950566.950595

Gelbal, S., & Kelecioğlu, H. (2007). Teachers' proficiency perceptions of about the measurement and evaluation techniques and the problems they confront. *Hacettepe University Journal of Education*, *33*(33), 135-145. https://dergipark.org.tr/tr/pub/hunefd/issue/7805/102347

Genç, M., & Çakmak Tolan, Ö. (2021). Play therapy practices in psychological and developmental disorders that are common in preschool period. *Current Approaches in Psychiatry*, *13*(2), 207-231. https://doi.org/10.18863/pgy.757366

Gomez, M. J., Ruipérez-Valiente, J. A., & Clemente, F. J. G. (2023). A systematic literature review of game-based assessment studies: Trends and challenges. *IEEE Transactions on Learning Technologies*, *16*(4), 500-515. https://doi.org/10.1109/TLT.2022.3226661

Gözüm, A. İ. C., & Kandır, A. (2020). Examination of the playfulness tendency and concentration levels of preschool children according to the time spent playing dijital games. *Journal of Kazım Karabekir Education Faculty*, *41*, 82-100. https://doi.org/10.33418/ataunikkefd.777424

Gündüz, M., Aktepe, V., Uzunoğlu, H., & Gündüz, D. D. (2017). The values obtained through the educative games to preschool children. *Muğla Sıtkı Koçman University Journal of Education*, *4*(1), 62-70. https://doi.org/10.21666/muefd.303856

Gür, H., & Kobak Demir, M. (2016). The views of the teacher candidates about the project of game based mathematics learning laboratory. *Necatibey Faculty of Education Electronic Journal of Science and Mathematics Education, 10*(1), 415-438. https://doi.org/10.17522/nefefmed.33834

Harlen, W. (2007). *The quality of learning: Assessment alternatives for primary education*. (Primary Review Research Survey 3/4), Cambridge. https://cprtrust.org.uk/wp-content/uploads/2014/06/research-survey-3-4.pdf

Hotaman, D. (2020). The importance of formative evaluation for success of online education. *The Journal of International Social Research*, *13*(73), 729-738. https://www.sosyalarastirmalar.com/articles/onlne-etmn-baarisi-aisindan-bmlendrc-deerlendrmenn-nem-the-importance-of-formative-evaluation-for-success-of-online-educ.pdf

Işıkoğlu Erdoğan, N., & Canbeldek, M. (2017). Assessment and evaluation in early childhood education. *Abant İzzet Baysal University Journal of Faculty of Education*, *17*(3), 1306-1327. https://doi.org/10.17240/aibuefd.2017.17.31178-338827

Karakuş, F. (2010). Teachers' views related to the alternative assessment methods in the program of secondary school education mathematics studies. *The Journal of Turkish Educational Sciences, 8*(2), 457-488. https://dergipark.org.tr/tr/pub/tebd/issue/26104/275043

Karasar, N. (2005). *Bilimsel araştırma yöntemi: Kavramlar ilkeler teknikler* (15th ed.). Nobel.

Karayol, S., & Temel, Z. F. (2018). An examination of 5-year-old children's problem-solving skills through play-based activities. *Cumhuriyet International Journal of Education*, *7*(2), 143-174. https://doi.org/10.30703/cije.393747

Katmada, A., Mavridis, A., & Tsiatsos, T. (2014). Implementing a game for supporting learning in mathematics. *Electronic Journal of e-Learning, 12*(3), 230-242. https://eric.ed.gov/?id=EJ1035662

Ketamo, H., & Devlin, K. (2014). Replacing PISA with global game based assessment. In C. Busch (Ed.), *Proceedings of the 8th european conference on games based learning Vol. 1* (pp. 258-264). Academic Conferences International Limited. https://web.stanford.edu/~kdevlin/Papers/Devlin-Ketamo_ECGBL2014.pdf

Kiili, K., & Ketamo, H. (2018). Evaluating cognitive and affective outcomes of a digital game-based math test. *IEEE transactions on learning technologies*, *11*(2), 255-263. https://doi.org/10.1109/TLT.2017.2687458

Kutlu, Ö., Doğan, C. D., & Karakaya, İ. (2010). *Ölçme ve değerlendirme: Performansa dayalı durum belirleme*. Pegem.

Leong, Y. X. E., & Toh, T. L. (2021). Game based assessment in mathematics classroom. *International Journal on Teaching and Learning Mathematics, 4*(1), 26-35. https://doi.org/10.18860/ijtlm.v4i1.10251

Looney, J. (2009). Assessment and innovation in education. *OECD Education Working Papers, 24*. OECD. https://doi.org/10.1787/222814543073

Mavridis, A., & Tsiatsos, T. (2017). Game-based assessment: Investigating the impact on test anxiety and exam performance. *Journal of Computer Assisted Learning, 33*(2), 137-150. https://doi.org/10.1111/jcal.12170

MoNE. (2023). *Millî eğitim bakanlığı ölçme ve değerlendirme yönetmeliği*. https://www.resmigazete.gov.tr/eskiler/2023/09/20230909-2.htm

MoNE. (2024a). *Primary school mathematics curriculum (Grades 1, 2, 3 and 4)*. https://tymm.meb.gov.tr/ogretim-programlari/ilkokul-matematik-dersi

MoNE. (2024b). *İlkokul matematik 1. sınıf ders kitabı*. Ministry of National Education.

Offenholley, K. H. (2012). Gaming your mathematics course: The theory and practice of games for learning. *Journal of Humanistic Mathematics, 2*(2), 79-92. https://scholarship.claremont.edu/jhm/vol2/iss2/7

Okatan, Ö., & Tagay, Ö. (2021). Pre-school teachers' views on play as a method of recognition and assessment of children. *Inonu University Journal of the Faculty of Education, 22*(3), 2140-2164. https://doi.org/10.17679/inuefd.892681

Olson, J. C. (2007). Developing students' mathematical reasoning through games. *Teaching Children Mathematics, 13*(9), 464-471. https://doi.org/10.5951/TCM.13.9.0464

Organisation for Economic Cooperation and Development [OECD] (2019). *PISA 2018 results (Volume I): What students know and can do*. OECD. https://doi.org/10.1787/5f07c754-en

Öğülmüş, S. (2019). Content analysis. *Ankara University Journal of Faculty of Educational Sciences (JFES), 24*(1), 213-228. https://doi.org/10.1501/Egifak_0000000729

Özata, M., & Coşkuntuncel, O. (2019). Opinions of secondary school mathematics teachers on th use of educational math games in mathematics teaching. *Mersin University Journal of the Faculty of Education, 15*(3), 662-683. https://doi.org/10.17860/mersinefd.619983

Özenç, M., & Çakır, M. (2015). E.xploring primary school teachers' competencies of alternative assessment and evaluation. *Elementary Education Online, 14*(3), 914-933. http://dx.doi.org/10.17051/io.2015.22900

Papert, S. (1993). *The children's machine: Rethinking school in the age of the computer*. Basic.

Patton, M. Q. (2018). *Nitel araştırma ve değerlendirme yöntemleri* [Qualitative research and evaluation methods] (M. Bütün, S. B. Demir, Trans. Eds.; 3rd ed.). Pegem.

Pellegrino, J. W., Chudowsky, N., & Glaser, R. E. (2001). *Knowing what students know: The science and design of educational assessment*. National Academies Press.

Pilten, P., Pilten, G., Divrik, R., & Divrik, F. (2017). Evaluation of mathematical game design skills of pre-service classroom teachers. *International Electronic Journal of Elementary Education, 10*(2), 255-264. https://doi.org/10.26822/iejee.2017236120

Pintér, K. (2010). Creating games from mathematical problems. *Primus, 21*(1), 73-90. https://doi.org/10.1080/10511970902889919

Russo, J. A., Bragg, L. A., & Russo, T. (2021). How primary teachers use games to support their teaching of mathematics. *International Electronic Journal of Elementary Education. 13*(4), 407-419. https://doi.org/10.26822/iejee.2021.200

Shepard, L. A. (2000). *The role of classroom assessment in teaching and learning* (CSE Technical Report 517). Los Angeles: University of California, Los Angeles, National Center for Research on Evaluation, Standards, and Student Testing (CRESST). https://jan.ucc.nau.edu/jf36/465word/shepardarticle.pdf

Skillen, J., Ricken, G., & Seitz-Stein, K. (2023). Game-based assessment of early mathematical competencies. A pilot study. *Empirische Sonderpädagogik, 15*(1), 38-60. https://doi.org/10.25656/01:27183

Şahin, E. (2019). *The effect of mind games on problem solving skills and tendencies of 4th grade primary school students* (Publication No. 580451) [Master's thesis, Balıkesir University]. Council of Higher Education Thesis Center.

Şanlıdağ, M., & Aykaç, N. (2021). The effect of mind games on students' attitudes of solving mathematical problems and reflective thinking skills for solving mathematical problems. *Muğla Sıtkı Koçman University Journal of Education, 8*(2), 597-611. https://doi.org/10.21666/muefd.846312

Şentürk, C. (2020). The effects of game based science learning experiences on academic achievement, retention, attitudes and learning process. *Milli Eğitim Dergisi, 49*(227), 159-194. https://dergipark.org.tr/tr/pub/milliegitim/issue/56322/616595

Tayfur, A. (2019). *Application of game supported assessment process on the teaching of 'the systems in our body' unit* (Publication No. 573842) [Master's thesis, Trabzon University]. Council of Higher Education Thesis Center.

Tekin, E. G. (2010). *Effect of formative assessment in mathematics education* ((Publication No. 279887) [Master's thesis, Marmara University]. Council of Higher Education Thesis Center.

Turgut, M. F., & Baykul, Y. (2019). *Eğitimde ölçme ve değerlendirme yöntemleri*. Pegem.

Türnüklü, D. A. (2000). Eğitimbilim araştırmalarında etkin olarak kullanılabilecek nitel bir araştırma tekniği: Görüşme. *Educational Administration: Theory and Practice, 24*(24), 543-559. https://dergipark.org.tr/en/pub/kuey/issue/10372/126941

Yazıcıoğlu, S., & Çavuş-Güngören, S. (2019). Investigate the effect of game-based activities on middle school students' science learning according to achievement, motivation, attitude and gender variables. *Necatibey Faculty of Education Electronic Journal of Science and Mathematics Education*, *13*(1), 389-413. https://doi.org/10.17522/balikesirnef.584673

Yıldırım, A., & Şimşek, H. (2021). *Sosyal bilimlerde nitel araştırma yöntemleri* (12th ed.). Seçkin.

Zengin, Y., Bars, M., & Şimşek, Ö. (2017). Investigation of using kahoot! and plickers in formative evaluation process in mathematics teaching. *Ege Journal of Education, 18*(2), 602-626. https://doi.org/10.12984/egeefd.318647

Necatibey Faculty of Education Electronic Journal of Science and Mathematics Education
Vol. 18, Issue 2, December 2024, pp. 460-479. ISSN: 1307-6086

Research Article

# Mathematics in Environmental Education: An Investigation of Preservice Teachers' Skills in Drawing and Interpreting Population Size Graphs

**Veli ÜNLÜ[1], Zeynep KOYUNLU ÜNLÜ[2]**

[1] Fatma Temel Turhan Science and Art Center, Türkiye, unluveli42@gmail.com, http://orcid.org/0000-0001-6055-1187

[2] Yozgat Bozok University, Faculty of Education, Türkiye, zeynepko.unlu@yobu.edu.tr, http://orcid.org/0000-0003-3627-1809

Received : 04.10.2024 Accepted : 22.11.2024

Doi: https://doi.org/10.17522/balikesirnef.1558829

---

*Abstract* – Drawing and interpretation graphs, as key mathematical skills, are widely used in teaching various subjects within science education. Population ecology, a topic in the environmental education course within the elementary education curriculum of Türkiye, is one such subject. In the context of this environmental education course, graphs are employed to teach and interpret the factors influencing the growth and decline of populations. This qualitative study aimed to examine the graph drawing and interpretation skills of preservice elementary school teachers in relation to population size. Fifty-seven preservice elementary school teachers from the elementary education department of a university in the Central Anatolia region of Türkiye participated in the study. These preservice teachers were provided with growth rate vs. time graphs of different populations and were asked to draw and interpret graphs representing change numbers of individuals vs. time. The findings revealed that many preservice teachers struggled with both drawing and interpreting population graphs. Specifically, most participants encountered difficulties while drawing and interpreting linear and particularly curvilinear graphs, as they failed to account for the simultaneous changes in two variables. The study's findings are expected to raise awareness about the importance of mathematical skills and the need for interdisciplinary collaboration in environmental education, as well as providing direction for future research.

*Keywords*: Environmental education, mathematics education, graph drawing skills, graph interpretation skills, population ecology, preservice teachers.

------------

Corresponding author: Veli ÜNLÜ, unluveli42@gmail.com

## Introduction

The teaching of many subjects in science requires the use of mathematics (Ministry of National Education [MoNE], 2018). Graph drawing and interpretation are among the essential mathematical skills used in science teaching and are crucial for science literacy, which involves interpreting and understanding scientific data (Osborne & Allchin, 2024). Graphs, as visual representations of quantitative and qualitative data, depict relationships between variables and enable comparisons by visualizing the data. Additionally, graphs help in summarizing, organizing, interpreting, and presenting data (Ateş et al., 2019). In this context, drawing graphs that align with specific research purposes based on observation and measurement results, as well as using tools such as frequency distributions, bar graphs, tables, and physical models, are integral parts of the science curricula currently and previously implemented in Türkiye (MoNE, 2005; MoNE, 2013; MoNE, 2018; MoNE, 2024). In current teaching practices within the updated science curriculum, students are often asked to create graphs based on data or interpret existing ones (MoNE, 2024).

Graph drawing and interpretation can be classified as mathematical and logical skills that contribute to the development of scientific processes, reasoning, visual literacy, and scientific reasoning skills, as noted in the literature (Ateş et al., 2019; Coştu et al., 2017; Krell et al., 2020). Since graph creation involves considering how two variables change together, covariational thinking/reasoning skills have been highlighted in recent studies (Altındiş et al., 2024; Basu & Panorkou, 2019; González, 2021, 2024). Covariational thinking is a complex cognitive process that requires understanding how two quantities change simultaneously (Carlson et al., 2002). In international exams, covariational reasoning skills involving graph drawing and interpretation are necessary to answer certain mathematics and science questions (Gant et al., 2023). Although research on covariational reasoning has traditionally focused on mathematics education, recent studies have extended that research into science education (Altındiş et al., 2024; González, 2021, 2024). Carlson et al. (2002) categorized covariational reasoning into five mental action (MA) levels. At the lowest level, MA1, students typically coordinate changes in one variable with changes in another. At the MA2 level, students recognize the direction of change; at MA3, they understand how much the dependent variable changes in relation to the independent variable; at MA4, they grasp the rate of change concerning the independent variable; and at the highest level, MA5, students comprehend how the instantaneous rate of change varies with consistent changes in the independent variable.

Many studies have been conducted to assess the graph reading and interpretation skills of students at various educational levels (Aydan & Dönel Akgül, 2021; Aydın & Tarakçı, 2018; Coştu et al., 2017; Erbilgin et al., 2015; Sezek, 2022; Sülün & Kozcu, 2005; Şahinkaya & Aladağ, 2013). Research has shown that students at the secondary school level (Erbilgin et al., 2015; Sülün & Kozcu, 2005; Tairab & Khalaf Al-Naqbi, 2004) and even at the university level (Aydan & Dönel Akgül, 2021; Aydın & Tarakçı, 2018; Coştu et al., 2017; Şahinkaya & Aladağ, 2013) encounter difficulties in drawing and interpreting graphs, skills that should be developed beginning in primary school. In the study by Coştu et al. (2017), it was found that preservice science teachers were more proficient in reading graphs in chemistry than interpreting them. Other researchers (Aydan & Dönel Akgül, 2021; Tairab & Khalaf Al-Naqbi, 2004; Taşar et al., 2002) have also indicated that students struggle more with drawing graphs than with reading or interpreting them. In a graphic drawing, difficulties may be experienced in issues such as determining and naming the axes, assigning numerical values to the axes, scaling and naming the graphic, determining the intersection points, starting the graphic axis from the appropriate place and continuing it (Aydan & Dönel Akgül, 2021; Aydın & Tarakçı, 2018). Research suggests that graph construction should be mastered before learning to read and interpret graphs (Ateş et al., 2019). Domain knowledge plays a critical role in interpreting scientific graphs, as the use of quantities during the interpretation process helps to form mental images and facilitates the development of covariational reasoning skills, particularly in cases of complex topics (Altındiş et al., 2024).

A literature review revealed that a limited number of studies have focused on the integration of mathematical skills in environmental education courses (Altındiş et al., 2024; Aydan & Dönel Akgül, 2021; Basu & Panorkou, 2019; González, 2021, 2024; Mumu et al., 2021; Özdemir, 2021; Sülün & Kozcu, 2005). For instance, in the study by Altındiş et al. (2024), university students were provided with two graphs representing the growth of two populations over time (exponential and logistic growth) and shown a video illustrating how the number of individuals changed over time. After watching the video, students were asked to explain the relationships depicted in each graph, compare the exponential and logistic graphs, and draw graphs showing how the number of individuals changed over time for each scenario. The results indicated that students with prior knowledge of graphing made more accurate predictions and interpretations. Another experimental study focused on 6th-grade students. It found that using simulations of the greenhouse effect was beneficial for exploring covariational relationships and developing complex reasoning (Basu & Panorkou, 2019).

González (2021, 2024) investigated the development of preservice mathematics teachers' covariational reasoning skills while modeling the relationship between carbon dioxide levels and global warming. The findings showed that as preservice teachers completed the tasks, their covariational reasoning improved, and using these skills enhanced their understanding and modeling of climate change. These results suggest that science topics can be better understood and explained when supported by mathematical skills. In a study by Mumu et al. (2021), middle school students' abilities to solve mathematical problems related to environmental education were examined. Solving problems about toxic waste, clean water, and flooding required knowledge of decimals, exponents, and fractions. Students' responses were categorized into four groups: (1) students who could solve mathematical problems and had environmental awareness, (2) students who could solve mathematical problems but had no interest in environmental issues, (3) students who could not solve mathematical problems but were interested in environmental issues, and (4) students who could not solve mathematical problems and were indifferent to environmental issues. The study found that the largest group consisted of students in the fourth category. In Özdemir's (2021) study, activities were developed that integrated environmental and mathematics education for 5th-grade students and the impact of those activities on students' views of sustainability was examined. The findings showed that combining mathematical and environmental concepts deepened students' superficial knowledge.

Population size, a topic covered in the environmental education course of the classroom teaching program, requires the use of covariational thinking and mathematical skills such as performing calculations and interpreting tables and graphs. Population size refers to the number of individuals that make up a population at a given time, and it is influenced by several factors, including birth rate, death rate, and migration. The population growth rate is determined by the sum of births and in-migration minus deaths and out-migration. In a population with a constant growth rate, the number of individuals increases steadily. In a population with a zero growth rate, the population size remains stable. If the growth rate increases smoothly over time, the population will experience an accelerating increase in the number of individuals per unit of time. Conversely, if the growth rate decreases linearly over time, the population size-and thus the rate of increase in the number of individuals-will decline per unit of time.

The development of preservice teachers' covariational reasoning skills related to graph drawing and interpretation is crucial, both for them to become scientifically literate individuals and to enhance the quality of their future classroom applications in this area. Integrating mathematics into environmental education not only fosters a deeper understanding of the subject matter but also improves mathematical skills and raises environmental awareness. As noted earlier, the existing literature highlights the need for more research on graph drawing and interpretation, often referred to as covariational reasoning, within environmental education. In light of this gap, the present study aimed to examine the population graph drawing and interpretation skills of preservice elementary school teachers. Specifically, the study sought to answer the following question: "*How proficient are preservice teachers in drawing and interpreting individuals vs. time graphs for populations based on given growth rate vs. time graphs?*"

## Method

### Research Model

In this study, preservice elementary school teachers were provided with growth rate-time graphs of different populations and asked to draw and interpret corresponding graphs of the number of individuals vs. time. The findings obtained from the qualitative data were interpreted. Therefore, this research was conducted using a basic qualitative approach (Merriam, 2013).

### Participants

The study involved 57 first-year preservice teachers enrolled in the elementary education department at a university situated in a small city center in Türkiye's Central Anatolia region.  The average age of the participants was 18 years. These preservice teachers were taking courses related to science and mathematics at the university, such as Basic Mathematics in Elementary School, Basic Science in Elementary School, and Environmental Education.

### Data Collection Tool and Process

A form prepared by the researchers was used as the data collection tool. Data were collected at the end of the courses in the spring semester of 2024. The form presented the participating preservice teachers with growth rate-time graphs for four distinct populations and asked them to draw and interpret the corresponding graphs of number of individuals vs. time. It was initially explained to the participants that the populations did not start with zero individuals. The first graph depicted a constant growth rate, the second a zero growth rate, the third a linearly increasing growth rate, and the fourth a linearly decreasing growth rate. To ensure the validity of the form, it was reviewed by three field experts, one language expert, and one expert in measurement and evaluation.

## Data Analysis

In the analysis process, an existing conceptual framework for classifying covariational thinking (e.g., Carlson et al., 2002; Thompson & Carlson, 2017) was not used. Instead, the preservice teachers' ability to draw and interpret graphs of number of individuals vs. time based on the given population size graphs was examined. During the data analysis, patterns were identified in the participants' responses (Merriam, 2013). In this process, data were coded for each graph given to the preservice teachers and categories were established. The types of graphs drawn and interpreted based on the given graphs were accepted as themes. Finally, the findings were interpreted.

The forms were coded by assigning numbers to each participant (PsT1, PsT2, ..., PsT57). Two researchers, one being an expert in science education and the other in mathematics education, independently evaluated the graphs drawn and the explanations provided by the preservice teachers. For each graph and its corresponding explanation, the researchers recorded whether the response was correct or incorrect and noted the reasoning behind their evaluation in a table. Any discrepancies between their evaluations were discussed until a consensus was reached, ensuring reliability. To enhance the credibility of the research, examples of the graphs drawn by the participants and their explanations are presented in the findings. The graphical drawings and explanations made by the preservice teachers regarding the given growth rate-time graph were analyzed independently of each other.

# Findings

In this study, preservice teachers were provided with four population growth rate-time graphs. The data obtained from their drawings (D) and explanations (E) were organized based on those four graphs and findings are presented below under the corresponding headings.

## Findings Related to the Drawing and Explanation of the First Graph

For the first question, preservice teachers were provided a graph showing a constant growth rate and were asked to draw a corresponding graph of number of individuals vs. time and explain their drawing. The findings obtained from the drawings and explanations for the first question are presented in Table 1.

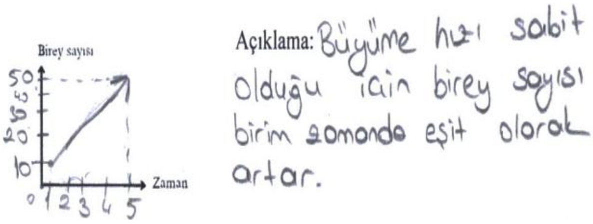
**Table 1** Findings Regarding Drawings and Explanations of Number of Individuals vs. Time Graph for the Graph with Constant Growth Rate*

| D/E | Categories | Frequency | Codes |
|---|---|---|---|
| Drawings | Correct | 6 | A positively sloped linear graph drawn starting from the y-axis (6) |
| | Incorrect | 46 | A graph drawn parallel to the x-axis (23)<br>A positively sloped linear graph drawn from the origin (16)<br>A graph that is not linear (4)<br>A graph drawn parallel to the y-axis (1)<br>A negatively sloped linear graph (1)<br>A graph for multiple time periods (1) |
| | No drawing | 5 | No drawing |
| Explanations | Correct | 26 | It was stated that the number of individuals would increase steadily over time (25)<br>It was explained that the number of individuals should initially be at a certain value and increase steadily over time (1) |
| | Incorrect | 30 | The number of individuals is balanced or constant (24)<br>Type 1 growth (3)<br>The number of individuals decreases (2)<br>Type 2 growth (1) |
| | No explanation | 1 | No explanation |

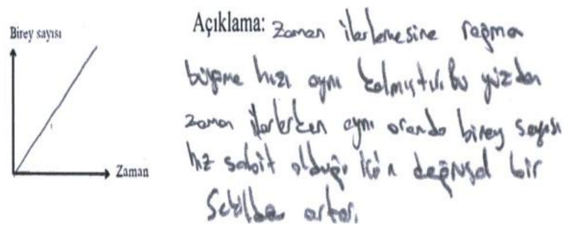
* Graphical drawings and explanations were analyzed independently of each other.

As can be seen in Table 1, 6 of the preservice teachers drew the graph of number of individuals vs. time correctly for the graph with a constant growth rate. The participants who drew it correctly depicted the graph as a positively sloped linear graph starting from the y-axis. A total of 46 participants produced incorrect drawings. Their inaccuracies were caused by drawing the graph parallel to the x-axis (n=23), starting it from the origin (n=16), not drawing it as a linear graph (n=4), drawing it parallel to the y-axis (n=1), drawing it as a negatively sloped linear graph (n=1), or combining different graphs (n=1). Five participants did not produce any drawings. Figure 1 shows examples of the drawings and explanations provided by these preservice teachers. Of the explanations given for the graphs, 26 were correct while 30 were incorrect. One participant did not provide an explanation.

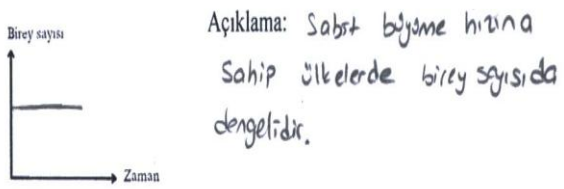
As shown in Figure 1, PsT10 provided both a correct drawing and explanation. PsT10 drew the graph as linear with a positive slope starting above the +y-axis. PsT25, however, drew the graph linearly with a positive slope but started from the origin, making PsT25's answer incorrect. The participants were informed that the population did not start with zero individuals. Therefore, the drawings of those who produced a linear graph with a positive slope starting from the origin, like PsT25, were marked incorrect, and this mistake was made by 16 of the preservice teachers, as seen in Table 1. Additionally, it was observed that PsT10 carefully considered the variables of population size and time while drawing the graph. Most participants who gave incorrect explanations believed that the population size remained in equilibrium or was constant (n=24). They assumed that with a constant growth rate, the population size would stay the same, attributing this to equal birth and death rates, which influenced both their drawings and explanations. The drawing and explanation provided by PsT21 in Figure 1 illustrate this common misunderstanding. Some participants also confused this graph with survival curves, interpreting it as a type 1 (n=3) or type 2 (n=1) survival curve. For instance, PsT49 stated that the population size would not be significantly affected by environmental factors, linking this to a type 1 survival curve in the provided explanation and drawing. In contrast, two participants thought that the population size would decrease in a population with a constant growth rate. PsT6's drawing and explanation exemplify this mistake, as PsT6 assumed that since the growth rate was constant, no new individuals were added, leading to a decrease in the population size.



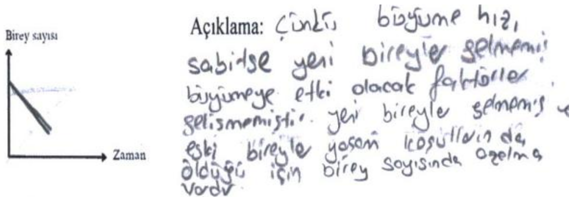

*" Since the growth rate is constant, the number of individuals increases equally in unit time." (PsT10)*



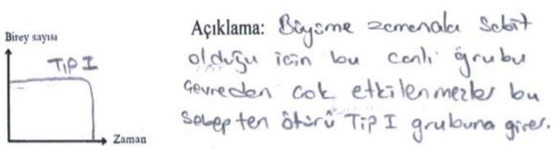

*"Despite the progression of time, the growth rate remains the same. Therefore, as time progresses, the number of individuals at the same rate increases linearly because the rate is constant." (PsT25)*




*"In countries with a stable growth rate, the number of individuals is also balanced." (PsT21)*




*"Because if the growth rate is constant, new individuals have not arrived and the factors that will affect growth have not developed. There is a decrease in the number of individuals because new individuals have not arrived and old individuals have died in living conditions." (PsT6)*




*" Since growth is constant over time, this group of organisms are not affected much by the environment, so they are included in the Type 1 growth." (PsT49)*

**Figure 1** Examples of Drawings of Preservice Teachers Who Correctly and Incorrectly Drew Graphs of the Number of Individuals vs. Time for a Population with Constant Growth Rate

## Findings Related to the Drawing and Explanation of the Second Graph

For the second question, the preservice teachers were given a population graph with a zero growth rate and were asked to draw and explain the related graph for number of individuals vs. time. The findings obtained from the preservice teachers' answers to the second question are presented in Table 2.

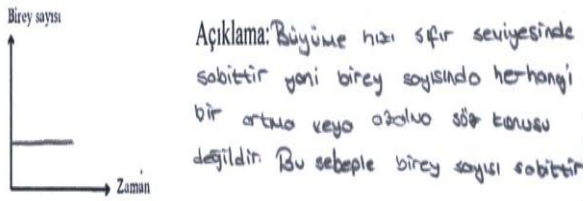
**Table 2** Findings Regarding Drawings and Explanations of Number of Individuals vs. Time Graph for the Graph With Zero Growth Rate*

| D/E | Categories | Frequency | Codes |
|---|---|---|---|
| Drawings | Correct | 22 | Graph drawn parallel to the x-axis (22) |
| | Incorrect | 31 | Graph coinciding with the x-axis (17)<br>Linear graph with negative slope (8)<br>Nonlinear graph (4)<br>Linear graph with positive slope (1)<br>Graph with multiple time intervals (1) |
| | No drawing | 4 | No drawing |
| Explanations | Correct | 26 | The number of individuals remains constant (26) |
| | Incorrect | 28 | The number of individuals decreases over time (13)<br>The number of individuals is zero because the growth rate is zero (6)<br>J-type growth (4)<br>No individuals are born (1)<br>The number of individuals increases over time (1)<br>Type 2 growth (1)<br>Type 3 growth (1)<br>Explanation unrelated to the number of individuals (1) |
| | No explanation | 3 | No explanation |

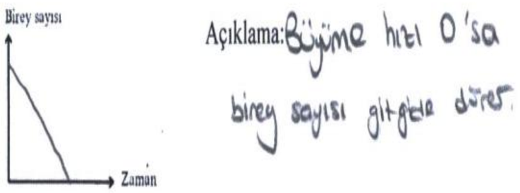
*Graphical drawings and explanations were analyzed independently of each other.

Twenty-two of the participants correctly drew the graph of number of individuals vs. time as a line parallel to the x-axis based on the graph with zero growth rate. The majority of those who drew the graph incorrectly assumed that the number of individuals would also be zero, reflecting this in their graphs (n=17). Some participants drew linear graphs with a negative slope, indicating a decrease in the number of individuals over time, since they interpreted a zero growth rate as causing a decline in population (n=8). Additionally, the responses of participants who did not draw a linear graph for the number of individuals vs. time (n=4), who drew a linear graph with a positive slope (n=1), or who combined different graphs (n=1) were also deemed incorrect. Based on the graph with zero growth rate, 26 of the preservice teachers drew the graph of number of individuals vs. time graph correctly, indicating that the number of individuals would remain constant over time.

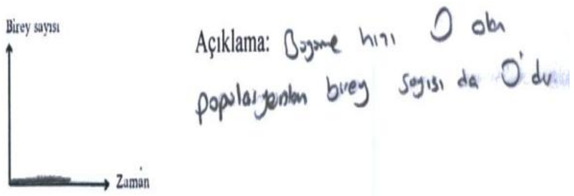
Figure 2 shows the answer of PsT16, one of the participants who produced both a correct drawing and explanation. PsT16 explained that the number of individuals would remain constant over time and drew the graph parallel to the x-axis. Most participants who provided incorrect explanations stated that the number of individuals would decrease over time (n=13). Figure 2 includes PsT20's drawing and explanation as an example. Some participants, such as PsT30, thought that the number of individuals would be zero because the growth rate was zero (n=6), thus drawing the graph of number of individuals vs. time to coincide with the x-axis (n=17). It was also observed that some participants whose answers were considered incorrect made explanations referencing type 2 and type 3 growth curves, as seen in survival curves of populations, like PsT49, or referenced J-type growth, a pattern seen in populations with exponential growth. Three participants did not provide any explanation. Interestingly, 4 participants drew the graph incorrectly but provided a correct explanation for the zero growth rate.



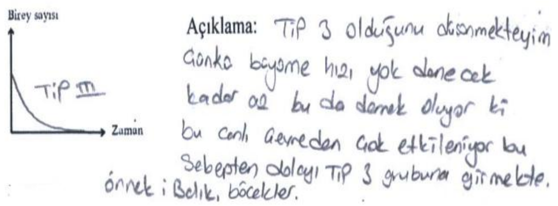

*" The growth rate is constant at zero level, that is, there is no increase or decrease in the number of individuals. Therefore, the number of individuals is constant." (PsT16)*




*"If the growth rate is zero, the number of individuals gradually decreases." (PsT20)*




*"If the population growth rate is 0, the number of individuals will also be 0." (PsT30)*




*" I think it is Type 3. Because its growth rate is almost non-existent. This means that this creature is greatly affected by the environment. For this reason, it falls into the type 3 group. Example: Fish, insects." (PsT49)*

**Figure 2** Examples of Drawings of Preservice Teachers who Correctly and Incorrectly Drew Graphs of the Number of Individuals vs. Time for a Population with Zero Growth Rate

## Findings Related to the Drawing and Explanation of the Third Graph

For the third question, the preservice teachers were given a graph where the growth rate increased linearly over time and were asked to draw and explain the graph of number of individuals vs. time accordingly. The findings are shown in Table 3.

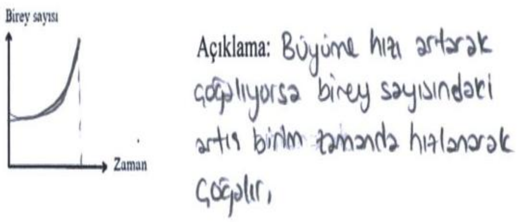
**Table 3** Findings Regarding Drawings and Explanations of Number of Individuals vs. Time Graph for the Graph with Linearly Increasing Growth Rate*

| D/E | Categories | Frequency | Codes |
|---|---|---|---|
| Drawings | Correct | 5 | Exponentially drawn graph starting from the +y-axis (5) |
| | Incorrect | 51 | Linearly drawn graph with positive slope (36)<br>Graph covering multiple time periods (9)<br>Linearly drawn graph with negative slope (3)<br>Exponentially drawn graph starting from the origin (2)<br>Graph parallel to the x-axis (1) |
| | No drawing | 1 | No drawing made |
| Explanations | Correct | 10 | The number of individuals increases more over time (10) |
| | Partially correct | 32 | The number of individuals increases over time (32) |
| | Incorrect | 14 | S-shaped growth (4)<br>The number of individuals decreases (2)<br>The number of individuals increases, remains constant, then decreases (1)<br>The number of individuals is in equilibrium or remains constant (1)<br>The number of individuals increases then remains constant (1)<br>The number of individuals increases logarithmically (1) |

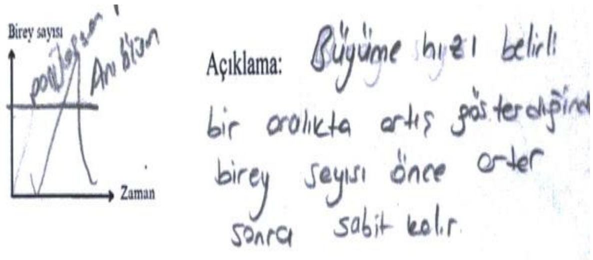
*Graphical drawings and explanations were analyzed independently of each other.

As can be seen in Table 3, only 5 of the preservice teachers drew the graph correctly. These participants represented the graph exponentially on the +y-axis. It was observed that the majority of participants (n=51) incorrectly drew the graph of number of individuals vs. time based on the graph with linearly increasing growth rate. Most of those who drew it incorrectly (n=36) represented both the growth rate and the number of individuals as increasing linearly over time. The answers of participants who combined different graphs (n=9), drew a linear graph with a negative slope (n=3), drew an exponential graph starting from the origin (n=2), or drew a graph parallel to the x-axis (n=1) were also considered incorrect. One participant did not make a drawing.

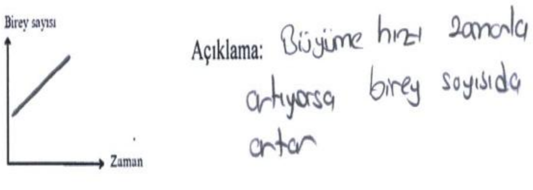
As seen in Table 3, 10 participants provided a correct explanation, stating that the number of individuals would increase exponentially over time. PsT52's response in Figure 3 is an example of a correct drawing and explanation. PsT52 drew the graph of number of individuals vs. time for the population with a linearly increasing growth rate as being exponential on the +y-axis, explaining that the number of individuals would increase more per unit of time. Responses of the participants that were considered partially correct (n=32) included accurate observations about the rise in the number of individuals over time. However, these increases were depicted as linear rather than exponential in their drawings. The explanations were generally similar to PsT3's response, shown in Figure 3, wherein the increase in the number of individuals was mentioned, but not in detail. A small number of participants (n=5) produced both correct drawings and comments. However, 14 participants provided incorrect explanations. Some explained and drew the number of individuals as decreasing when the growth rate increased over time, possibly because they were thinking about the later stages of S-shaped growth curves. Additionally, 5 participants made correct explanations but drew their graphs incorrectly, and 1 did not provide any explanation.



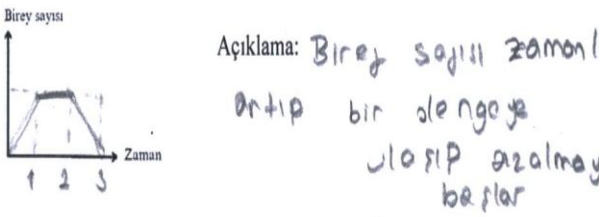

*"If the growth rate increases, the number of individuals increases at an accelerating rate per unit time." (PsT52)*




*"Since the growth rate increases within a certain range, the number of individuals first increases and then remains constant." (PsT44)*




*"If the growth rate increases over time, the number of individuals also increases" (PsT3)*




*"The number of individuals increases over time, reaches a balance and begins to decrease" (PsT24)*

**Figure 3** Examples of Drawings of Preservice Teachers who Correctly and Incorrectly Drew Graphs of the Number of Individuals vs. Time for a Population whose Growth rate Increases Linearly over Time

## Findings Related to the Drawing and Explanation of the Fourth Graph

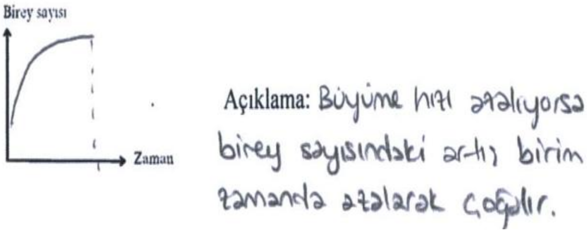
For the fourth question, the participating preservice teachers were asked to draw and explain the graph of number of individuals vs. time graph for a graph whose growth rate decreases linearly over time. The findings obtained from the participants' answers to the fourth question are shown in Table 4.

**Table 4** Findings Regarding Drawings and Explanations of Number of Individuals vs. Time Graph for a Population whose Growth Rate Decreases Linearly Over Time*

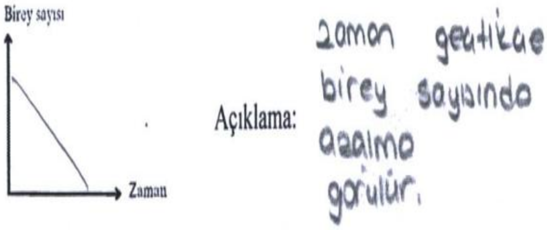
| D/E | Categories | Frequency | Codes |
|---|---|---|---|
| Drawings | Correct | 1 | Logarithmic graph drawn starting from the +y-axis (1) |
| | Incorrect | 55 | Linear graph with a negative slope (22)<br>Graph with multiple time intervals (13)<br>Linear graph with a positive slope (9)<br>Non-logarithmic graph starting from the +y-axis (8)<br>Logarithmic graph starting from the origin (2)<br>Graph parallel to the x-axis (1) |
| | No drawing | 1 | No drawing |
| Explanations | Correct | 12 | The number of individuals increases more slowly (12) |
| | Partially correct | 2 | The number of individuals increases over time (2) |
| | Incorrect | 39 | The number of individuals decreases (30)<br>Number remains stable or constant (2)<br>Sudden deaths occur (1)<br>S-type growth (1)<br>Decreases, then remains stable (1)<br>Type 3 growth (1)<br>Increases, then decreases (1)<br>Increases very rapidly (1)<br>Explanation unrelated to the number of individuals (1) |

*Graphical drawings and explanations were analyzed independently of each other.

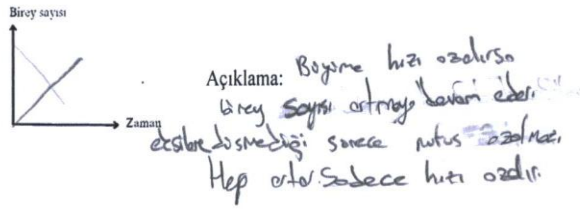
As seen in Table 4, only one participant produced a correct drawing. The majority of participants who produced incorrect drawings (n=22) drew the graph of number of individuals vs. time based on a decreasing linear growth rate as a linearly decreasing graph. Additionally, some participants combined different graphs in their drawings (n=13). As can be seen in Table 4, only 12 of the participants stated that the number of individuals would increase but the increase would occur at a decreasing rate. Some participants who correctly stated that the number of individuals would decrease over time drew this decrease linearly rather than logarithmically (n=11). Only one participant (n=1) produced both a correct drawing and explanation. Two participants whose answers were considered partially correct stated that the number of individuals would decrease over time and did not elaborate on the nature of the decrease. Most of the participants whose answers were deemed incorrect explained that the number of individuals decreased (n=30). Additionally, 11 participants drew the graph incorrectly but provided a correct explanation. Some participants produced explanations and drawings suggesting that the number of individuals would increase as the growth rate decreased over time. Furthermore, some interpreted the graph as a J-shaped graph. Figure 4 provides examples of the drawings and explanations of the participants who correctly and incorrectly drew the graph of number of individuals vs. time for a population whose growth rate decreases linearly over time.



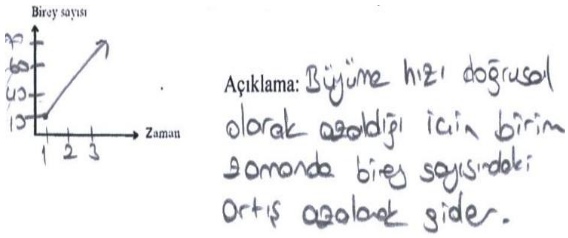

*"If the growth rate decreases, the increase in the number of individuals increases by decreasing in unit time." (PsT52)*




*"Over time, the number of individuals decreases." (PsT43)*




*"If the growth rate decreases, the number of individuals continues to increase. The population does not decrease unless it is negative. It always increases. It just decreases in speed." (PsT25)*




*"Since the growth rate decreases linearly, the increase in the number of individuals per unit time decreases." (PsT10)*

**Figure 4** Examples of Drawings of Preservice Teachers who Correctly and Incorrectly Drew Graphs of the Number of Individuals vs. Time for a Population whose Growth Rate Decreases Linearly Over Time

Only PsT52 stated that the number of individuals in a population with a linearly decreasing growth rate increases at a decreasing rate over time and made a correct drawing. Most of the participating preservice teachers, like PsT43, thought that the number of individuals would decrease over time and drew a linear graph with a negative slope. From the drawings and explanations of PsT25 and PsT10, it can be understood that some preservice teachers predicted that the number of individuals would increase at a decreasing rate over time; they made correct explanations but could not accurately transfer those explanations to the graph.

## Conclusions, Discussion, and Suggestions

In this study, preservice primary school teachers were given growth rate-time graphs of different populations and asked to draw and interpret the corresponding graphs for number of individuals vs. time. The results showed that these preservice teachers had difficulty in drawing and interpreting population graphs and particularly in interpreting linear and especially curvilinear graphs. Only one participant drew and interpreted all the graphs correctly. This participant is a graduate of a science high school where science and mathematics courses are predominant. In this respect, it can be said that this participant has a basic background. The most common mistakes can be listed as follows: (1) Starting from the origin when drawing the graph of the number of individuals vs. time, even though they were initially told that the number of individuals in the population is not zero; (2) believing that the number of individuals in a population with a constant growth rate will also remain constant; (3) thinking that the number of individuals in a population with a zero growth rate will also be zero; (4) assuming that the number of individuals in a population with a linearly increasing growth rate will also increase linearly over time; and (5) believing that the number of individuals in a population with a smoothly decreasing growth rate will also decrease over time. This may be due to the fact that the preservice teachers did not sufficiently understand and inquiry the graph, and lacked prior knowledge (Shah & Hoeffner, 2002). It is thought that the most common mistakes in graph drawing are due to the fact that the preservice teachers do not reflect the relationships between the data on the graph, but rather think of this relationship as in the first graph and transfer it to the second graph as a picture. This reveals that they have a misconception of graphs as pictures (Roth & Bowen, 2001). Also these results indicate that most of the preservice teachers could not account for the simultaneous change of two variables. They had more difficulty drawing curvilinear graphs than linear graphs. This may be due to the fact that preservice teachers tend to create linear graphs because they do not evaluate the relationship in the graph by not looking at the whole graph (Leinhardt vd., 1990). Similarly, some research results related to physics showed that students struggled to interpret rates of change in nonlinear or curvilinear graphs, which was associated with deficiencies in their mathematical knowledge (McDermott et al., 1987; Planinic et al., 2013). Furthermore, in a study conducted with calculus students, it was observed that the students had difficulty creating images of the rate of change and could not accurately represent or interpret the increasing and decreasing rates of functions (Carlson et al., 2002). In addition, as a result of this research, the difficulty experienced by preservice teachers in determining the starting point of the graph is similar to the study of Aydın and Tarakçı (2018).

In the present study, some of the participating preservice teachers interpreted the graphs given to them correctly, but they could not transfer their correct interpretations into the drawing and interpretation of a new graph. This was particularly evident in the responses to the first and second questions, which required drawing and interpreting linear graphs. From this perspective, it can be inferred that the majority of participants remained at the comprehension/understanding level in graph drawing and interpretation; they had not progressed to higher levels such as application, analysis, and synthesis. These results align with findings from several previous studies on graph interpretation and drawing. For instance, Tairab and Khalaf Al-Naqbi (2004) and Taşar et al. (2002) found that students had more difficulty drawing graphs than reading and interpreting them. In a study on preservice science teachers, it was determined that, similar to the results of this research, the preservice teachers' levels of reading and interpreting graphs were better than drawing graphs in most cases (Aydan & Dönel Akgül, 2021).

In light of the findings of this study, it may be beneficial to support preservice teachers in developing the ability to draw and interpret graphs through modeling. In this process, the initial number of individuals should be determined, tables should be created for the number of individuals at specific time intervals, and graphical drawings and interpretations should be made based on the created tables. Using numbers during the process of interpreting and creating graphs can help provide a more concrete understanding and promote correct thinking (Altındiş et al., 2024). As indicated in the findings of this study, some preservice teachers drew graphs of the number of individuals vs. time by assigning values to the growth-rate time graph. Additionally, there should be an emphasis on questioning the operations performed. Otherwise, the interpretation of graphs that are consistently explained and observed may be incorrect, resulting in a lack of skill development in this area. Furthermore, the results obtained from this study highlight the necessity of integrating different disciplines. Difficulties in graph interpretation and drawing may hinder the understanding of concepts that are intended to be taught (Coştu et al., 2017). In this context, it is essential to know and utilize the mathematical skills, methods, and techniques specific to mathematics education. When needed, lecturers working in mathematics and science disciplines can work together and support teaching practices, thus enabling interdisciplinary collaboration. Another suggestion based on the results of this research is that the environmental education course should be designed and implemented to provide not only knowledge but also practical skills. These skills could include not only the ability to draw and interpret graphs, as evaluated in this study, but also probabilistic thinking and proportional reasoning (Lawson et al., 2000). Some subjects may inherently be more suitable for teaching certain skills. Therefore, it is crucial to first determine which skills can be developed within the context of the environmental education course and to design appropriate teaching environments accordingly. Future research may focus on the design, implementation, and evaluation of a skills-based environmental education course.

**Compliance with Ethical Standards**

*Disclosure of potential conflicts of interest*

The authors declare that they have no competing interests.

*Funding*

No funding was received for conducting this study.

*CRediT author statement*

The article was collaboratively written by two authors, with each contributing equally to its content.

*Research involving Human Participants and/or Animals*

The study involves human participants. Ethics committee permission was obtained from Yozgat Bozok University, Social and Human Sciences Ethics Committee.

## Çevre Eğitiminde Matematik: Öğretmen Adaylarının Popülasyon Büyüklüğü Grafiği Çizme ve Yorumlama Becerilerinin İncelenmesi

**Özet:**

Matematiksel bir beceri olan grafik çizme ve yorumlama, fen eğitiminde pek çok konunun öğretiminde kullanılmaktadır. Sınıf öğretmenliği programında yer alan çevre eğitimi dersindeki popülasyon ekolojisi bu konulardan biridir. Çevre eğitimi dersi kapsamında bir popülasyonun büyüme ve küçülme nedenlerinin öğretimi ve yorumlanmasında grafiklere başvurulmaktadır. Nitel olarak yürütülen bu çalışmada sınıf öğretmen adaylarının popülasyon grafiği çizme ve yorumlama becerilerinin incelenmesi amaçlanmıştır. Araştırmaya Türkiye'nin İç Anadolu Bölgesinde yer alan bir üniversitenin sınıf eğitimi bölümünde öğrenim gören 57 öğretmen adayı katılmıştır. Sınıf öğretmen adaylarına farklı popülasyonlara ait büyüme hızı-zaman grafikleri verilmiş, onlardan birey sayısındaki değişimleri zamana göre gösteren grafikler çizmeleri ve yorumlamaları istenmiştir. Sonuçlar öğretmen adaylarının popülasyon grafiği çizme ve yorumlama konusunda güçlük yaşadığını göstermiştir. Doğrusal ve özellikle eğrisel grafiklerin çizim ve yorumlanmasında öğretmen adaylarının çoğu iki değişkenin eş zamanlı değişimini hesaba katamamıştır. Araştırma sonuçlarının çevre eğitiminde matematiksel becerilerin kullanımına ve disiplinler arası işbirliğinin gerekliliğine ilişkin farkındalığı artıracağı, yapılacak çalışmalara rehber olacağı düşünülmektedir.

*Anahtar kelimeler*: Çevre eğitimi, matematik eğitimi, grafik çizme becerisi, grafik yorumlama becerisi, popülasyon ekolojisi, öğretmen adayları.

## References

Altindis, N., Bowe, K. A., Couch, B., Bauer, C. F., & Aikens, M. L. (2024). Exploring the role of disciplinary knowledge in students' covariational reasoning during graphical interpretation. *International Journal of STEM Education, 11*(32), 1-23. https://doi.org/10.1186/s40594-024-00492-5

Ateş, S., Temiz, K. B., Yüksel, İ., Şahin F., İnaltun, H., & Yalvaç, E. (2019). *Scientific reasoning (logical thinking),* Palme.

Aydan, B., & Dönel Akgül, G. (2021). Pre-service science teachers' skills of reading, interpreting and drawing graphs. *ATLAS Journal International Refereed Journal on Social Sciences, 7*(40), 1768-1796. http://dx.doi.org/10.31568/atlas.678

Aydın, A., & Tarakçı, F. (2018). The investigation of the pre-service science teachers' abilities to read, interpret and draw graphs. *Elementary Education Online, 17*(1), 469-488. https://doi.org/10.17051/ilkonline.2018.413806

Basu, D., & Panorkou, N. (2019). Integrating covariational reasoning and technology into the teaching and learning of the greenhouse effect. *Journal of Mathematics Education, 12*(1), 6-23. https://doi.org/10.26711/007577152790035

Carlson, M., Jacobs, S., Coe, E., Larsen, S., & Hsu, E. (2002). Applying covariational reasoning while modeling dynamic events: A framework and a study. *Journal for Research in Mathematics Education, 33*(5), 352–378. https://doi.org/10.2307/4149958

Coştu, F., Ercan, O., & Coştu, B. (2017). Pre-service science teachers' levels of graph reading and interpretation. *Buca Faculty of Education Journal, 44*, 194-213. https://dergipark.org.tr/tr/pub/deubefd/issue/35768/401197

Erbilgin, E., Arıkan, S., & Yabanlı H. (2015). Assessing line graph comprehension and construction skills. *Ahi Evran University Journal of Kırşehir Education Faculty, 16*(2), 43-61. https://dergipark.org.tr/tr/pub/kefad/issue/59450/854099

Gant, A. L., Paoletti, T., & Corven, J. (2023). Exploring the prevalence of covariational reasoning across mathematics and science using TIMSS 2011 assessment items. *International Journal of Science and Mathematics Education, 21*, 2349-2373. https://doi.org/10.1007/s10763-023-10353-2

González, D. A. (2021). The progression of preservice teachers' covariational reasoning as they model global warming. *Journal of Mathematical Behavior, 62,* Article 100859. https://doi.org/10.1016/j.jmathb.2021.100859

González, D. A. (2024). The role of preservice teachers' quantitative and covariational reasoning in understanding climate change. *Cognition and Instruction*, *42*(2), 294-326. https://doi.org/10.1080/07370008.2024.2314499

Krell, M., Redman, C., Mathesius, S., Krüger, D. & van Driel, J. (2020). Assessing pre-service science teachers' scientific reasoning competencies. *Research in Science Education*, *50*, 2305-2329. https://doi.org/10.1007/s11165-018-9780-1

Lawson, E. A., Clark, B., Meldrum- Cramer, E., Falconer, K., A., Sequist, J. M., & Kwon, Y-J. (2000). Development of scientific reasoning in college biology: do two levels of general hypothesis-testing skills exist?. *Journal of Research in Science Teaching 37*(1), 81-101. https://doi.org/10.1002/(SICI)1098-2736(200001)37:1<81:AID-TEA6>3.0.CO;2-I

Leinhardt, G., Zaslavsky, O., & Stein, M. K. (1990). Functions, graphs, and graphing: tasks, learning and teaching. *Review of Educational Research*, *60*(1), 1-64. https://doi.org/10.3102/00346543060001001

McDermott, L. C., Rosenquist, M. L., & Van Zee, E. H. (1987). Student difficulties in connecting graphs and physics: Examples from kinematics. *American Journal of Physics*, *55*(6), 503-513. https://doi.org/10.1119/1.15104

MoNE. (2005). *İlköğretim fen ve teknoloji dersi (6, 7 ve 8. sınıflar) öğretim programı*.

MoNE. (2013). *İlköğretim kurumları (ilkokullar ve ortaokullar) fen bilimleri dersi (3, 4, 5, 6, 7 ve 8. sınıflar) öğretim programı*.

MoNE. (2018). *İlköğretim kurumları (ilkokullar ve ortaokullar) fen bilimleri dersi (3, 4, 5, 6, 7 ve 8. sınıflar) öğretim programı*.

MoNE. (2024). *Fen bilimleri dersi öğretim programı (3, 4, 5, 6, 7 ve 8. sınıflar) Türkiye Yüzyılı Maarif Modeli.*

Merriam, S. B. (2013). *Qualitative research a guide to design and implementation*. John Wiley & Sons.

Mumu, J., Prahmana, R. C. I., Sabariah, V., Tanujaya, B., Bawole, R., Warami, H., & Monim, H. O. L. (2021). Students' ability to solve mathematical problems in the context of environmental issues. *Mathematics Teaching Research Journal*, *13*(4), 99-111. https://files.eric.ed.gov/fulltext/EJ1332360.pdf

Osborne, J., & Allchin, D. (2024). Science literacy in the twenty-first century: informed trust and the competent outsider. *International Journal of Science Education*, 1–22. https://doi.org/10.1080/09500693.2024.2331980

Özdemir, D. (2021). Integration of mathematics and environmental education: change in the views of 5th grade students toward sustainability. *Acta Didactica Napocensia, 14*(2), 194-202. https://doi.org/10.24193/adn.14.2.14

Planinic, M., Ivanjek, L., Susac, A., & Milin-Sipus, Z. (2013). Comparison of university students' understanding of graphs in different contexts. *Physical Review Special Topics-Physics Education Research, 9*(2), Article 020103. https://doi.org/10.1103/PhysRevSTPER.9.020103

Roth, W. M. & Bowen, G. M. (2001). Professionals reads graphs: a semiotic analysis. *Journal for Research in Mathematics Education, 32*(2), 159-194. https://www.jstor.org/stable/pdf/749672.pdf

Sezek, F. (2022). Determination of graphic literacy levels of primary teacher candidates. *International Social, Mentality and Researcher Thinkers Journal, 8*(63), 1651-1660. https://smartofjournal.com/?mod=tammetin&makaleadi=&key=64390

Shah, P. & Hoeffner, J. (2002). Review of graph comprehension research: implications for instruction. *Educational Psychology Review, 14*, 47-69. https://doi.org/10.1023/A:1013180410169

Sülün, Y., & Kozcu, N. (2005). The mistakes of 8. class students in primary schools in perception on interpretation of graphic questions about the subject of environment and population in the high school entrance exams. *Erzincan University Journal of Education Faculty, 7*(1) 25-33. https://dergipark.org.tr/en/download/article-file/67179

Şahinkaya, N., & Aladağ, E. (2013). The views of pre-service class teachers on graphics. *Adiyaman University Journal of Social Sciences, 6*(15), 309-328. https://dergipark.org.tr/tr/download/article-file/15087

Tairab, H. H., & Khalaf Al-Naqbi, A.K. (2004). How do secondary school science students interpret and construct scientific graphs?. *Journal of Biological Education, 38*(3), 127-132. https://doi.org/10.1080/00219266.2004.9655920

Taşar, M. F., Kandil İngeç, Ş., & Ünlü Güneş, P. (2002, September 16-18). *Grafik çizme ve anlama becerisinin saptanması* [Conference presentation]. V. Ulusal Fen Bilimleri ve Matematik Eğitimi Kongresi, Ankara, Türkiye.

Thompson, P. W., & Carlson, M. P. (2017). *Variation, covariation, and functions: Foundational ways of thinking mathematically*. In J. Cai (Ed.), First compendium for research in mathematics education (pp. 421-456). Reston, VA: NCTM.

**Necatibey Faculty of Education Electronic Journal of Science and Mathematics Education**
**Vol. 18, Issue 2, December 2024, pp. 480-502. ISSN: 1307-6086**

**Research Article**

# Trending Themes on the Nature of Science in Science Education: A Bibliometric Analysis with VOSviewer

**Serpil KARA[1]**

[1] Necmettin Erbakan University, Ahmet Keleşoğlu Faculty of Education, Türkiye, serpilkara@erbakan.edu.tr, http://orcid.org/0000-0002-0482-7617

Received : 11.09.2024 Accepted : 26.11.2024

Doi: https://doi.org/10.17522/balikesirnef.1548266

---

*Abstract* – This study aimed to identify themes that indicate trends in the nature of science (NoS) in science education. The Web of Science (WoS) database was used for this purpose. A bibliometric approach was adopted, and the VOSviewer software was used to identify and systematically map research trends in the field of scientific innovation and to identify intellectual relationships in this network. In this context, the study covered the years 2013-2023. The categories “Education & Educational Research” and “Education Scientific Disciplines” were selected in the WoS database, and only article-type studies were included by excluding other publication categories obtained in the breakdown of the results. Finally, 263 articles were analyzed . According to the findings, the “Education Educational Research” WoS category is at the forefront. The publications made according to the years show alternating trends of increase and decrease. Lederman, N.G. is the author with the highest number of citations. Erduran, S. is the author with the highest number of studies and the highest total link strength. The majority of publications were in the journal “Research In Science Education.” The words “scientific literacy” and “history of science” were used most frequently after “nature of science” and “science education” among the keywords. The most frequently repeated words in the abstracts of the articles were "questionnaire" and "interview". Finally, the countries ranked in the top three in terms of total link strength and the most cited countries were the USA, England, and Türkiye, respectively.

*Keywords*: Nature of science, science education, bibliometric analysis, VOSviewer, history of science and scientific literacy

------------

Corresponding author: Serpil KARA, serpilkara@erbakan.edu.tr

## Introduction

The primary purpose of science is to explain natural phenomena in a logical and organized manner, develop theories, and discover basic principles and concepts. By integrating scientific methods into educational environments, the goal is for students to conduct research to understand the world and actively participate in the scientific process to understand how scientific knowledge and the nature of science (NoS) are formed (Ministry of National Education [MoNE], 2018).

The NoS is defined as a discipline that seeks answers to issues such as the working methods of scientists, the definition of science, the interaction between science and society, and the values and beliefs in scientific knowledge (Lederman 2007). The literature reveals that the NoS is handled on the basis of dimensions (Khishfe & Abd-El-Khalick, 2002; Lederman, 2007). These dimensions are considered the changeability, empiricism, theoretical basis, and imaginative and creative aspects of scientific knowledge, as well as the influence of the social and cultural environment, and the theoretical, observational, and inferential nature of scientific knowledge.

Khishfe and Abd-El-Khalick (2002) proposed three approaches for teaching NoS: historical, indirect, and direct reflective. On the other hand, the Family Resemblance Approach can be integrated into science curricula to enable students to learn and understand science holistically in its epistemic, cognitive, and social dimensions (Kaya & Erduran, 2016). Therefore, in the related literature, various approaches are used in the field of science education to teach the NoS (Gören & Kaya, 2023).

### NoS in Science Education

An important research area in science education is the NoS (Gören & Kaya, 2023), and one of the main goals of this education is to enable students to understand the NoS (AAAS, 1993; NRC, 1996; McComas, 2004; NGSS, 2013; Voss et al., 2023). Therefore, it is argued that curricula should be organized and implemented to develop an understanding of science and NoS (Abd-El-Khalick, F., & Lederman, N. G.,2000). When evaluated globally, science education reform documents and curricula include outcomes related to the development of NoS understanding (NGSS Lead States, 2013; NRC, 1996). In the context of science education, NoS plays a critical role in teaching students the basic principles, methods, and processes of science. One of the main objectives of our country's Science Curriculum, which aims to raise all individuals to be science-literate, is to help them understand how scientific knowledge is created by scientists, the processes through which this knowledge is created, and how it is used in new research (MoNE, 2018).

Effective teaching of the NoS is not only limited to the transfer of knowledge in science education (Çakıcı, 2009), but also aims to improve students' abilities to apply knowledge, generate new questions, and question existing knowledge (Ayvacı & Akdemir, 2017). In this context, teaching strategies and curricula must be designed to support students' scientific thinking and problem-solving skills. Such teaching helps students develop a deeper understanding of the NoS and gain perspectives on the social, ethical, and practical dimensions of science (Hwang et al., 2015). Indeed, understanding the NoS in science education not only increases the scientific knowledge of individuals, but it also develops social consciousness and responsibility (Crowther et al., 2005) and enables individuals to think critically by understanding the production of scientific knowledge (Jimenez-Aleixandre et al., 2000). Therefore, in the literature, both teachers (Adsiz & Kutluca, 2023; Kurt & Kaya, 2023) and students (Gülmez Güngörmez & Akgün, 2020; Ozan & Uluçınar Sağır, 2020; Stadermann & Goedhart, 2020; Yacoubian, 2021) identify or develop an understanding of the NoS in the context (Bugingo et al., 2024), or it is observed that various methods are used for effective teaching (Gören & Kaya, 2023).

The ESERA conference, organized by the European Science Education Research Association, encompasses worldwide research in the field of science education. When the studies on NoS in the ESERA-2009 conference were examined, it was found that the studies on NoS were directly related to science subjects: NoS and student-teacher views, conceptual understanding, the relationship between NoS and teaching methods, and even epistemological beliefs (Öztürk & Kaptan, 2014). Studies focusing on various components have been conducted to determine or develop NoS understanding in the global context. In this context, it is important to identify trends in the current literature on the Nature of Science (NoS) and teaching in the field of science education. One of the studies that has frequently been used recently to identify trends in terms of various components is bibliometric analysis.

## Bibliometric Analysis and Study Importance

Bibliometric analysis is preferred to obtain information about the breadth and quantity of the topic (Comarú et al., 2021), as well as for a general understanding of its nature and to guide future studies (Pradana et al., 2023). While it allows us to reveal the evolutionary nuances of a particular field, it also sheds light on emerging areas in that field. Bibliometric analysis involves the application of quantitative techniques to bibliometric data (Donthu et al., 2021).

When the related literature is examined, it is seen that there are studies on science education (Comarú et al., 2021; Effendi et al., 2021; Maryanti et al., 2023; Solehuddin et al., 2022; Sönmez & Hastürk, 2020; Tosun, 2024) and limited bibliometric studies on NoS (Kurtuluş & Bilen, 2021; Yanuarti & Suprapto, 2021). Kurtuluş and Bilen (2021) conducted a bibliometric analysis of studies on NoS in science education, published in science, physics, chemistry, and biology education journals indexed in the WoS database between 1986-2019. On the other hand, Yanuarti and Suprapto (2021) conducted a bibliometric analysis of studies on NoS in science education between 2011 and 2020 based on the Scopus database. However, studies on bibliometric analysis of NoS in science education remain limited. One of the important goals of science education is for students to understand the NoS (AAAS, 1993; NGSS Lead States, 2013; NRC, 1996), and goals related to the NoS are included in science education reform documents and curricula (AAAS, 1993; NGSS Lead States, 2013; NRC, 1996). On the other hand, since NoS teaching has critical importance in the dimension of raising scientifically literate individuals (Widowati et al., 2017), and since the understanding of NoS has long been among the goals of science education (Voss et al., 2023), it is important to know the trends of studies on NoS in science education.

It is thought that presenting the trends of the studies on the NoS in science teaching and the innovations at the current point through a bibliometric analysis will shed light on the researchers who will work in this field in terms of identifying the gaps, new scientific paths, and study topics, and will make a significant contribution to the establishment of scientific cooperation globally. This approach will assist in establishing scientific cooperation globally. The bibliometric analysis on science teaching aims to provide researchers with important information on existing literature and to help them determine which areas are more extensively studied and which topics are not sufficiently researched. This provides an opportunity to focus on previously unexplored topics and is expected to contribute to the identification of new research topics by addressing research gaps. The analysis of the existing findings will enable scientists to develop innovative research questions and new methods in the field. By examining collaborations between research conducted in different countries and institutions at the global level, bibliometric analysis provides a broader perspective on the research community and increases opportunities for international collaboration. As a result, this study is expected to serve as a guide for new and existing researchers and to make a significant contribution to the development of the field, by providing a roadmap for science teaching and NoS.

In this context, this study was designed to answer the following questions.

1. Regarding research on NoS in science education;

a. What is the WoS category distribution?

b. How is the distribution based on years?

c. What is the ranking of the top 10 authors with the most studies?

d. What is the distribution of journals in which their articles were published?

e. What is the citation analysis network of the authors?

f. What is the co-author analysis network?

g. What is the keyword analysis network?

h. What is the analysis network for the most frequently used words in the article abstracts?

ı. What is country citation analysis network?

## Method

The research design, data collection, and analysis processes are described below.

### Research Design

In this research, a bibliometric approach was adopted to identify and systematically map research trends in the field of scientific innovation (Prahani et al., 2024) and intellectual relationships in this network (Li & Xu, 2022). The VOSviewer program, a visualization tool used in the bibliometric approach, was applied. VOSviewer uses elements such as scientific publications, journals, researchers, research institutions, organizations, countries, keywords, and terms to create and visualize networks of relationships through links such as coauthorship, collaboration, citation, and cocitation (Aria & Cuccurullo, 2017; Van Eck & Waltman, 2022).

### Data Collection and Analysis

In this study, the widely used Web of Science (WoS) database (Sarkar et al., 2022), which hosts scientific documents in all disciplines, was used to obtain data. WoS, the first bibliometric database established by the Institute for Scientific Information (ISI) (Pranckute, 2021), was preferred because it is widely used for academic literature search and selection (Agrifoglio et al., 2021). It is a multidisciplinary and selective database consisting of various specialized indexes grouped according to the type or theme of the indexed content. It is a core collection of six major citation indexes (Pranckute, 2021).

On 21.07.2024, access to the WoS was provided through the university database affiliated with the researcher, and the contents indexed there were used as criteria. In the WoS database, a search using the keywords “nature of science” and “science education” under the title "topic" yielded 722 results. The nature of science has gained an important place in educational curricula in recent years because of its foundational impact and comprehensive scope. Prior to 2013, scientific studies generally showed an increasing trend. Between 2013 and 2023, scientific studies showed fluctuating trends, and the importance given to science in the education system increased, which makes it important to analyze these periods. Moreover, the fact that data will continue to flow as of 2024 provides an important context for understanding the trends and developments over the last decade. Therefore, analyzing studies from this decade is of primary importance, to better understand the changes in the nature and practices of science. Given that the NoS is a very broad and long-established subject, has recently gained an important place in education curricula, and will continue to have active data flow in 2024, the aim was to analyze studies from 2013 to 2023 to determine the trend over the last ten years. In this context, 502 publications were identified. In the next stage, the categories of “Education & Educational Research” and “Education Scientific Disciplines” were selected in the WoS database, and only articles as document type were analyzed. A total of 263 publications were identified and analyzed. Papers, book chapters, article reviews, early access, editorial material, and book reviews were not included in the dataset or evaluated. The selection process of the articles to be included in the study with the PRISMA flow diagram (Moher et al., 2009) is shown in Figure 1.

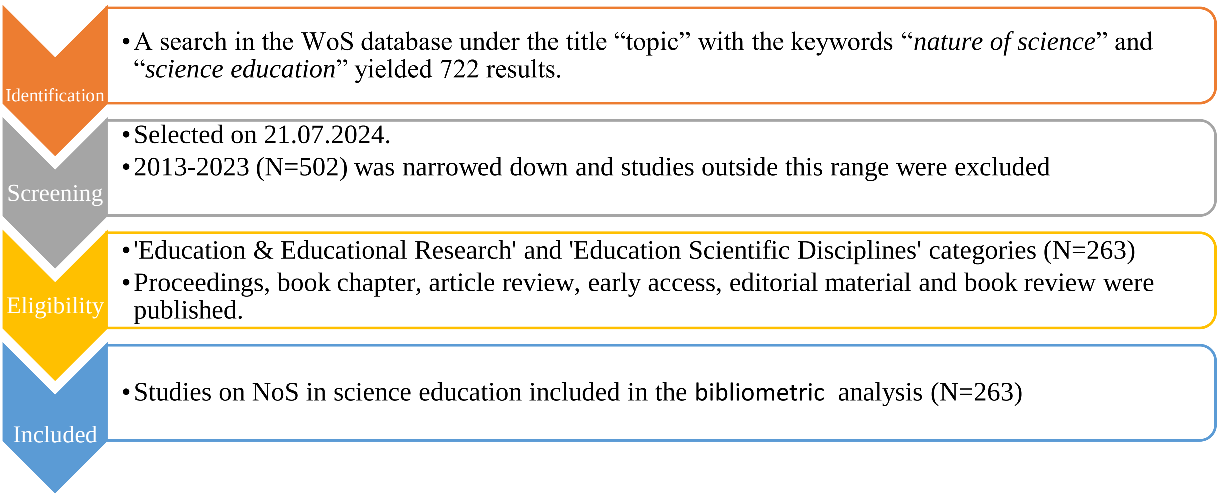



**Figure 1** Selection Process of the Articles to be Included in the Study

## Findings

### WoS Category Breakdown

Firstly, the distribution of the first ten categories reached in the research conducted using the keywords "nature of science" and "science education" is given in Table 1. The distribution of the categories of 'Education & Educational Research' and 'Education Scientific Disciplines' are given in Table 2.

**Table 1** WoS Category Distribution of Studies on Nos in Science Education

| Category | Number of articles |
|---|---|
| Education & educational research | 532 |
| History philosophy of science | 103 |
| Education scientific disciplines | 94 |
| Cultural studies | 21 |
| Biology | 15 |
| Geosciences multidisciplinary | 7 |
| Psychology educational | 6 |
| Chemistry multidisciplinary | 4 |
| Physics multidisciplinary | 4 |
| Psychology multidisciplinary | 4 |

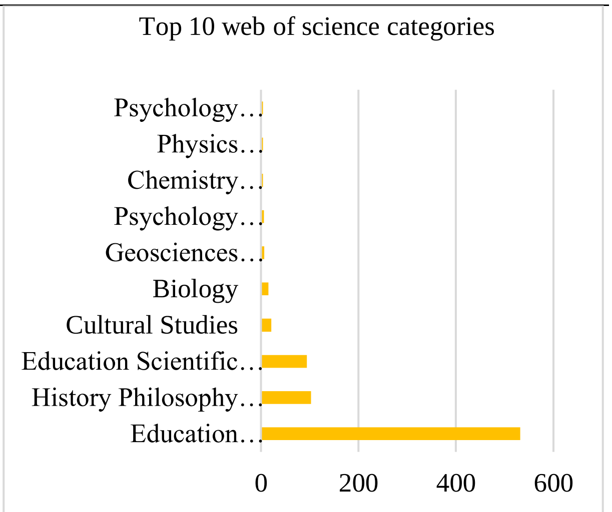
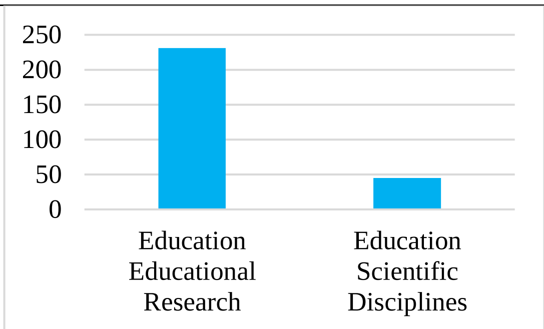



**Table 2** WoS Category Breakdowns Selected within the Scope of the Research

| Category | Number of articles |
|---|---|
| Education & educational research | 231 |
| Education scientific disciplines | 45 |

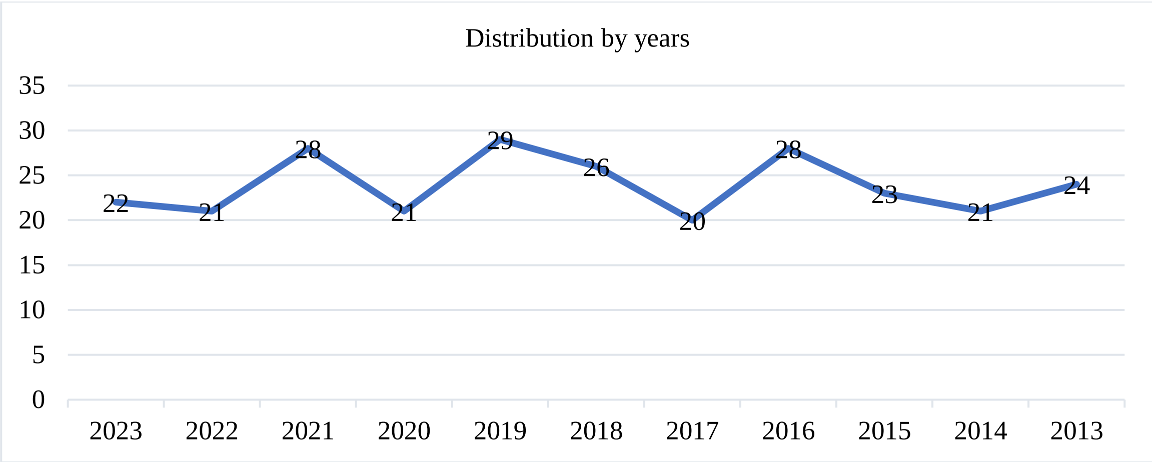



When Tables 1 and 2 are examined together, it can be seen that the "Education & Educational Research" category is at the forefront of the WoS category distribution of the studies on the NoS in science education.

### Distribution of Research on NoS in Science Education by Year

Table 3 shows a distribution of research conducted between 2013 and 2023.

**Table 3** Distribution of NoS Research in Science Education by Year



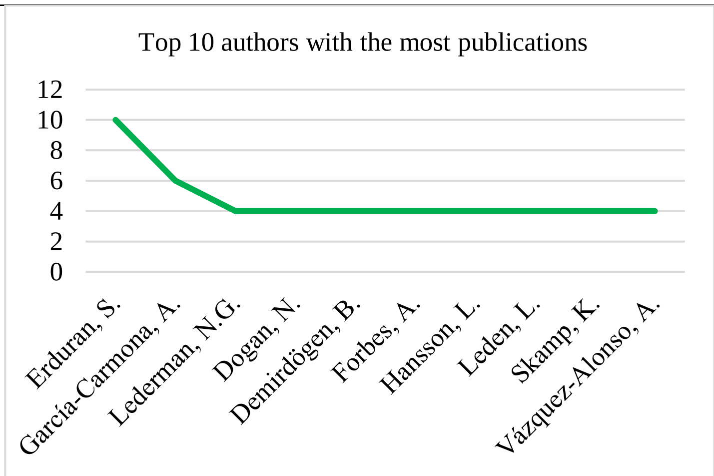

An analysis of the yearly distribution reveals that *f*=24 articles were published in 2013, *f*=21 in 2014, *f*=23 in 2015, *f*=28 in 2016, *f*=20 in 2017, *f*=26 in 2018, *f*=29 in 2019, *f*=21 in 2020, *f*=28 in 2021, *f*=21 in 2022 and *f*=22 in 2023. Considering the number of studies by years, it is seen that there was an increase in 2015, 2016, 2018, 2019 and 2021, and a decrease in 2017, 2020, and 2022. Starting from 2023, considering the increase in the number of studies, it can be said that studies on the NoS in science education will increase in the coming years 2024 and 2025.

## Findings of the Top 10 Authors with the Most Number of Studies

Table 4 lists the top 10 authors who conducted the most studies on NoS in science education between 2013 and 2023, as revealed in the analysis.

**Table 4** Top 10 Authors with the Most NoS Work in Science Education

| Author name | Number of articles |
|---|---|
| Erduran, S. | 10 |
| García-Carmona, A. | 6 |
| Lederman, N.G. | 4 |
| Dogan, N. | 4 |
| Demirdögen, B. | 4 |
| Forbes, A. | 4 |
| Hansson, L. | 4 |
| Leden, L. | 4 |
| Skamp, K. | 4 |
| Vázquez-Alonso, A. | 4 |



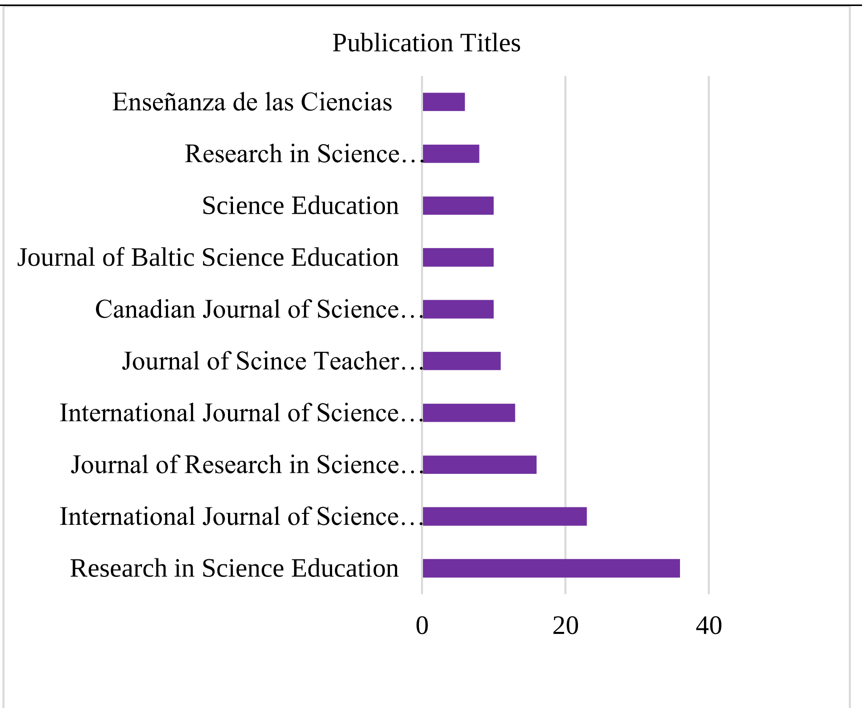
When Table 4 is examined, the author who conducted the most studies on the subject was Erduran, S.

## Distribution of journals in which research on NoS in science education is published

The findings regarding the distribution of the journals in which the studies were published are presented in Table 5.

**Table 5** Distribution of Journals in which Articles Related to the NoS in Science Education were Published

| Journals | Number of articles |
|---|---|
| Research in science education | 36 |
| International journal of science education | 23 |
| Journal of research in science teaching | 16 |
| International journal of science and mathematics education | 13 |
| Journal of science teacher education | 11 |
| Canadian journal of science, mathematics, and technology education | 10 |
| Journal of baltic science education | 10 |
| Science education | 10 |
| Research in science and technological education | 8 |
| Enseñanza de las ciencias | 6 |




According to Table 5, the studies on NoS in science education were mostly found in Research in Science Education and International Journal of Science Education.

## Citation Analysis Network of Authors in Research on the NoS in Science Education (Citation of Authors)

At least 2 publications and at least 2 different citation criteria were used in the bibliometric analysis to identify author networks. The results that met the criteria are visualized in Figure 2.

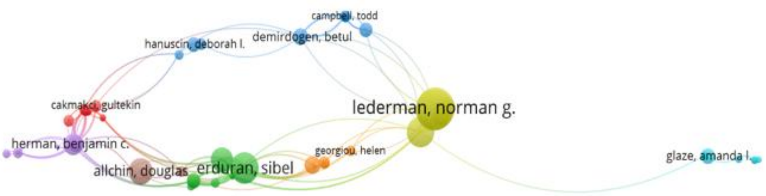



**Figure 2** Citation Analysis of the Authors

In the author citation network map in Figure 2, the most cited authors are Lederman, N. G. (365 citations, 5 publications), Lederman, J. S. (352 citations, 3 publications), Erduran, S. (201 citations, 10 publications), Schwartz, R. S. (198 citations, 2 publications), and Allchin, D. (146 citations, 3 publications). A total of 8 clusters, 194 links, and 286 total link strengths were determined over 59 units with links between authors.

**Analysis Network of Co-Authorship of Authors in Research on NoS in Science Education (Co-Authorship of Authors)**

A bibliometric network was created based on the criteria; at least 1 publication and at least 1 citation, to identify the most contributing, connected, and collaborating authors. The bibliometric network of the data is presented in Figure 3.

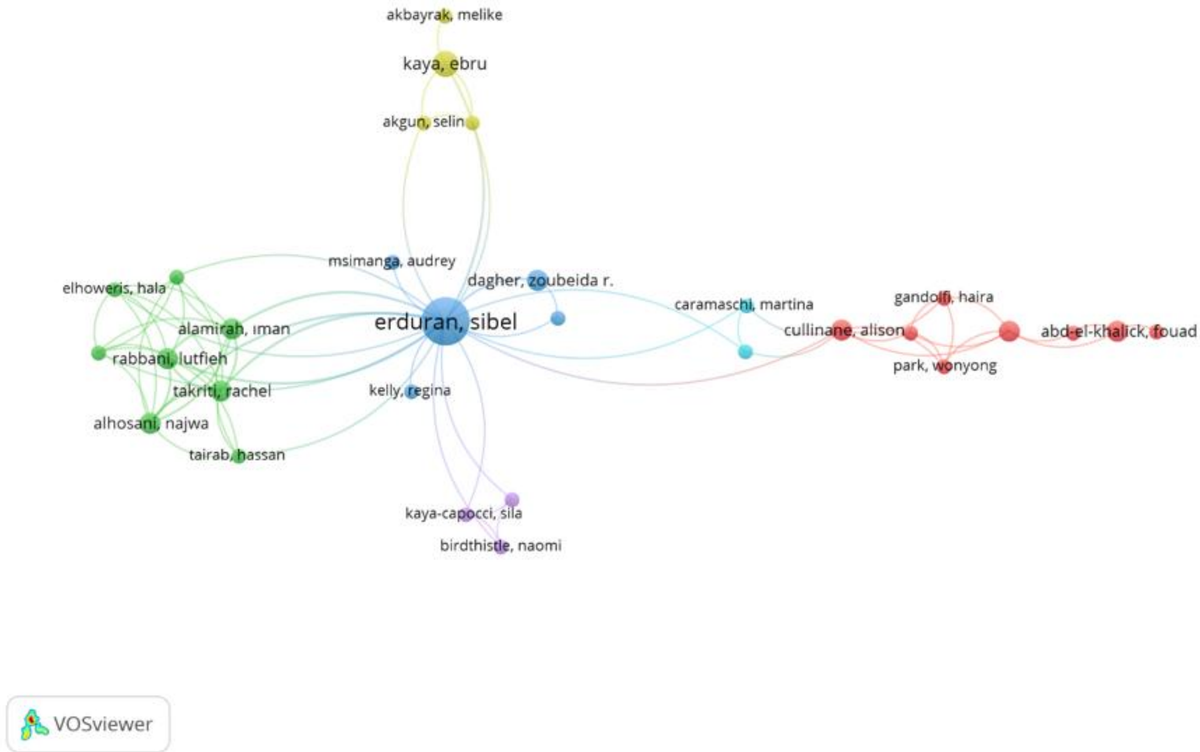



**Figure 3** Co-authorship of the Authors Analysis

According to Figure 3, the bibliometric network formed consisted of 71 links merging into 6 clusters with a total of 83 link strengths obtained. The authors with the highest total link strength are Erduran, S. (10 publications, 201 links, total link strength 27), Doğan, N. (4 publications, 32 links, total link strength 18), Çakmakçı, G. (4 publications, 33 links, total link strength 17), Yalaki, Y., and Irez, S. (3 publications, 28 links, total link strength 15), and Furman, M. (2 publications, 29 links, total link strength 13).

**Analysis Network of Keywords in Research on Nos in Science Education (Co-Occurrence of Author Keywords)**

To determine the analysis network of keywords in the studies on NoS in science education, the analysis type was selected as "Co-occurrence," and the analysis unit was selected as "Author Keywords" in the VOSviewer program. The network obtained as a result of the analysis is shown in Figure 4.

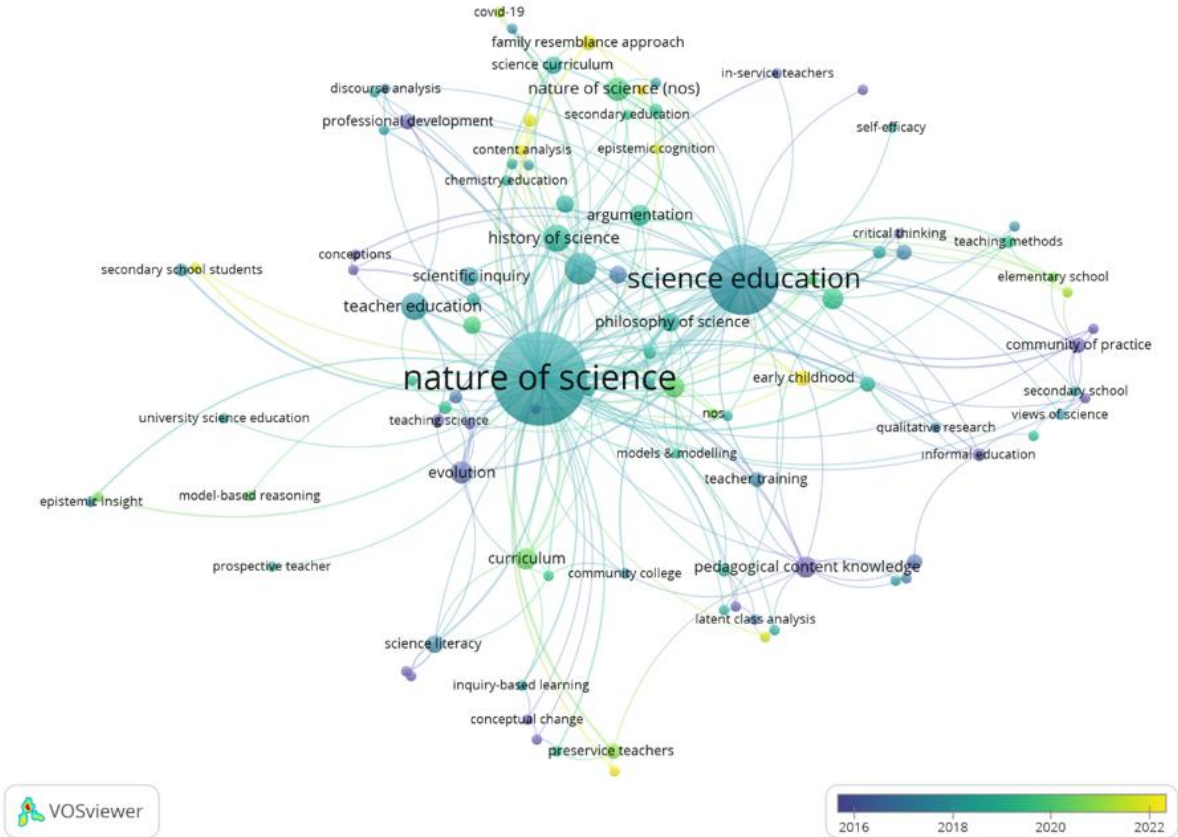



**Figure 4** Co-occurrence Analysis of the Author Keywords

According to Figure 4, 99 out of 687 words were found in the keyword analysis to meet this threshold of occurring at least 2 times in frequency. Thus, 20 clusters, 288 links, and 475 total link strengths were identified. Among the keywords in the published articles, "nature of science" was the most frequent (138), followed by "science education" (76), "scientific literacy" (16), "history of science" with 12 occurrences, and "teacher education" with 12 occurrences. As a result of the analysis of 209 related items, a total of 24 clusters, 790 links, and 1278 total link strengths were identified. Since the 2020s (Figure 4, yellow areas), keywords such as "family resemblance approach; early childhood; content analysis; special education" have begun to appear in the studies.

**Analysis Network of the Most Frequently Used Words in the Abstracts of Articles in Research on the NoS in Science Education**

"Abstract field" and "binary counting" methods were selected, and the total number of terms in the abstracts, of the publications was determined as 5889. The minimum number of repetitions was set to 10, and the analysis was performed on the 92 terms that met this criterion. The word network obtained as a result of the analysis is shown in Figure 5.

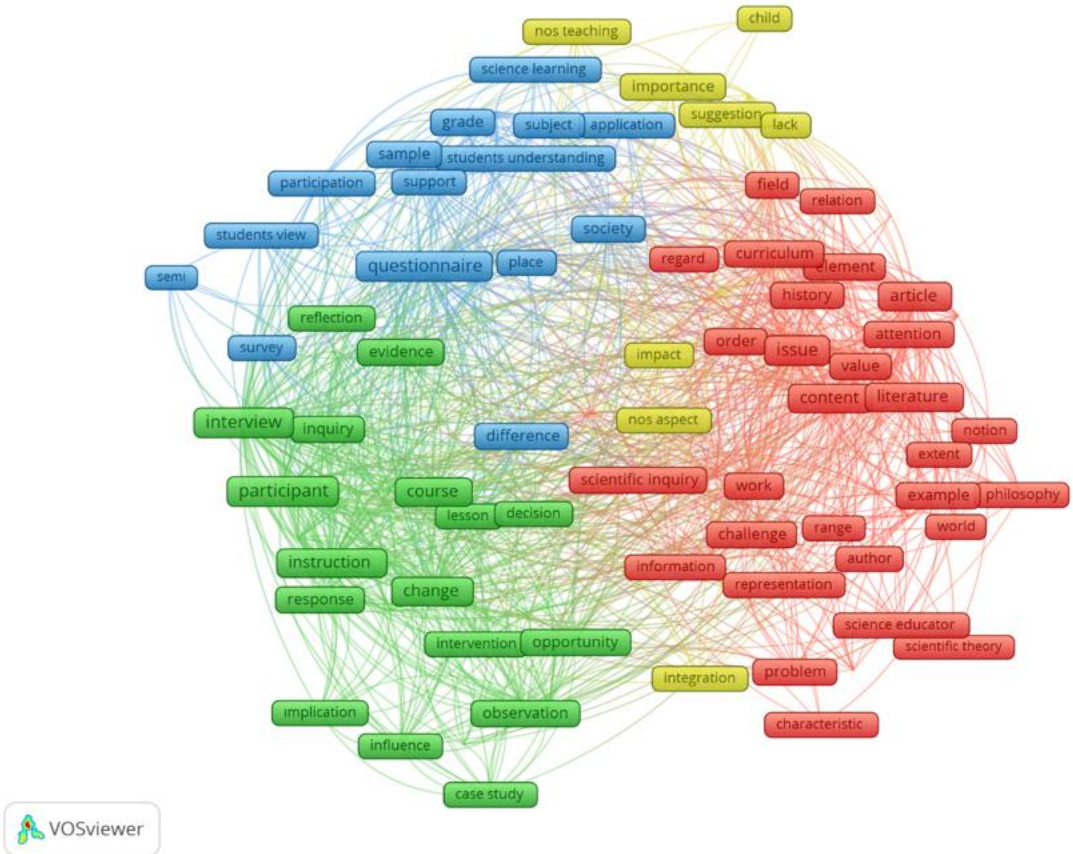



**Figure 5** Most Frequently Used Words in the Abstracts

The map revealed 4 thematic clusters [Cluster 1 red (*f*=37), Cluster 2 green (*f*=26), Cluster 3 blue (*f*=18), and Cluster 4 yellow (*f*=11)]. The largest circle in each cluster represents the dominant keyword. When Figure 5 is analyzed, the most frequently repeated words in the abstracts of the related articles are questionnaire (*f*=51, blue cluster,) interview (*f*=49, green cluster,) participant (*f*=49, green cluster,) issue (*f*=48, red cluster,) and course (*f*=40, green cluster).

**Citation Analysis Network of Countries in Research on The Nos in Science Education (Citation of Countries)**

To create a network map based on the citations received by the countries from which the publications originated, 31 observation units were analyzed in accordance with the condition that a specific country published at least two works and received at least two citations. The network map is shown in Figure 6.

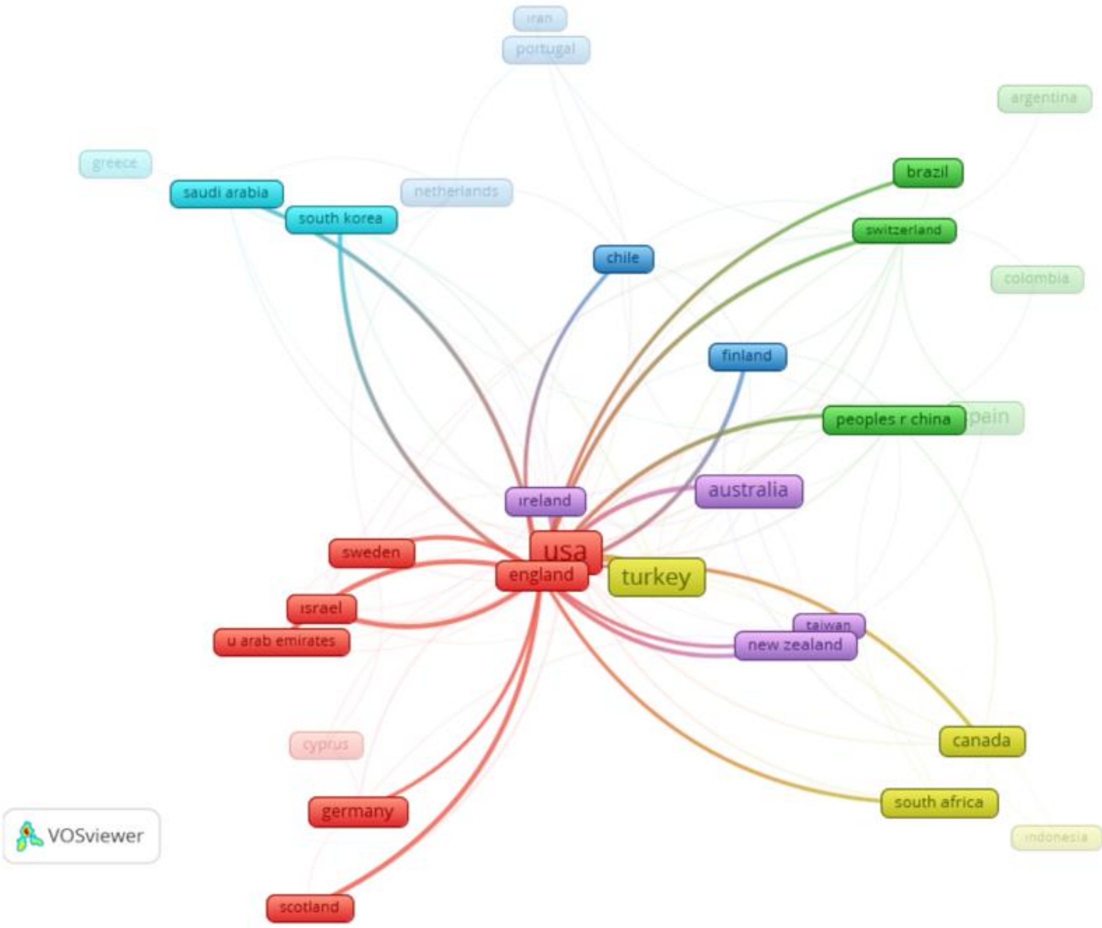



**Figure 6** Analysis of Country Citations

In the citation analysis network of countries, the top three countries in terms of total link strength and the most cited countries were the USA (2135 citations, 78 publications), England (606 citations, 15 publications), and Türkiye (735 citations, 41 publications). The level of relationships among the 31 units is explained by 7 clusters, 113 links, and 256 total link strengths.

## Conclusions, Discussions, and Suggestions

In this study, a bibliometric analysis of articles related to NoS studies in science education in the categories of "Education & Educational Research" and "Education Scientific Disciplines" in the WoS database was presented. The studies included in this context cover the years 2013-2023. According to the findings, most articles in the Web of Science category distribution were in the "Education & Educational Research" category.

When the findings regarding the distribution of the studies by year were evaluated, the intensity of the studies increased in 2016 ($f$=28), 2019 ($f$=29), and 2021 ($f$=28). It then started to decline in 2023 ($f$=22). However, in general, studies on NoS in science education fluctuate over time. There are concrete suggestions on how to make changes in science curricula to transform various aspects of the NoS into a more holistic and inclusive form (Caramaschi et al., 2022). From a global perspective, considering  the goals related to the NoS in science education reform documents and curricula (AAAS, 1993; NGSS Lead States, 2013; NRC, 1996) and similarly in our country in both the primary education curriculum (MoNE, 2018) and teacher training programs (CoHE, 2018), we expect that studies on the subject will gain intensity.

Erduran, S. and García-carmona, A. are the authors who conducted the most studies on the subject, with Research in Science Education being the journal where their work is most frequently published. On the other hand, the authors with the highest number of citations are Lederman, N. G., Lederman, Judith S., and Erduran, S. The authors with the highest total link strength are Erduran, S., Doğan, N., and Çakmakçı,. G. Tosun (2024) concluded that Lederman N.G. is one of the most cited authors in science education research, indicating that the subject of the NoS is intensively studied.

The most frequently repeated keywords in the published articles were "NoS," "science education," "scientific literacy," "history of science," and "teacher education." This can be interpreted as the studies mostly focused on teachers' scientific literacy, and the history of science conceptions. Tosun (2024) found that one of the frequently used keywords and core topics in the results of bibliometric analysis of science education research articles from the past 40 years was NoS and that NoS was among the most preferred topics between 2007 and 2021. In the current study, it was concluded that keywords such as familial approach, early childhood, content analysis, and special education started to be included in studies in the 2020s. As a result of a systematic review of studies on the family resemblance approach to the NoS in science education, it was determined that the number of studies using FRA has increased in the last decade (Cheung & Erduran, 2023), which supports the findings of the current study.

On the other hand, although there has been limited interest in teaching NoS in early childhood in the past years (Bell & Clair, 2015), a study based on the NoS approach in early childhood found that conversations about the NoS are possible for the youngest children, and that focusing on science at an early age will contribute to the NoS literature (Hansson et al., 2020). In recent years, there has been a focus on the importance of understanding the NoS in early childhood. It will represent one of the important research topics in this field. In recent years, content analysis studies have been conducted analyzing the NoS under various components (Bett et al., 2023; Cheung & Erduran, 2023; Jaenudin et al., 2021; Okan & Kaya, 2023; Suryani et al., 2022). the NoS is an important subject that can be examined under a wide variety of variables; therefore, it can guide different study topics. Considering that content analyses, like bibliometric analyses, are conducted to determine trends and tendencies related to any subject, it is important to repeat them at certain time intervals. Similarly, studies on the NoS in the field of special education have been included in recent years (Librea-Carden & Mulvey, 2023; Librea-Carden et al., 2021), and this field offers potential for further study on the NoS.

According to the findings related to the most frequently used words in the abstracts of the articles, the research priorities of the NoS researchers in science education were identified as questionnaire, interview, participant, issue, and course. In this case, the studies focused on obtaining teachers' opinions on NoS, derived from course and subject/activity practices. Finally, it was concluded that the countries in the top three in terms of total link strength in the citation analysis network of countries and the countries with the highest number of citations were the USA, the UK, and Türkiye. The finding that the USA and the UK are the leading countries in science education research (Tosun, 2024) supports the findings of this study. On the other hand, just as teachers' views on the NoS play a key role in how they teach science (Takriti, et al., 2024), it is also important to obtain their students' perspectives on the NoS as a result of activities and practices to understand and succeed in science (Karataş Öztürk et al., 2023; Yacoubian, 2021). It can be concluded that there has been an increase in questionnaire and interview studies on the NoS. After the review of the studies on

FRA in the NoS in science education, it was found that interviews and questionnaires were the leading methods (Cheung & Erduran, 2023). Considering that views and students are among the most used keywords in a bibliometric study on the NoS (Kurtuluş & Bilen, 2021), in another content analysis study, students were the sample in a significant portion of the studies (Taşkın, 2021), and recent studies have progressed on determining student views (Gülmez Güngörmez & Akgün, 2020; Ozan & Uluçınar Sağır, 2020; Stadermann & Goedhart, 2020), it can be considered that what is meant as participant in the results of the current study constitutes the group of students. However, because one of the most frequently repeated words among the keywords of the articles of the current study was "teacher education," the group characterized as "participants" in this study consists of teachers.

In future research, it should be considered that teachers should receive quality training to effectively transfer knowledge. In this context, a model in which different practices are implemented and evaluated in undergraduate teacher training programs (e.g., see Wahbeh & Abd-El-Khalick, 2014) or a model based on science teachers' pedagogical content knowledge resources can be used for more effective teaching of NoS. In addition, it may be recommended to conduct similar studies using domestic databases such as ULAKBİM and Council of Higher Education (CoHE) Thesis Archive, and internationally using databases such as Scopus and PubMed.

**Compliance with Ethical Standards**

*Disclosure of potential conflicts of interest*

I have no conflict of interest to declare.

*Funding*

The study was not funded

*CRediT author statement.*

All stages of the study were provided by the author

*Research involving humans and/or animals*

No data were collected from human participants in the study. The research is a document review. All data were obtained by accessing the Web of Science-WoS database in the e-Library documentation unit of Necmettin Erbakan University

## Fen Eğitiminde Bilimin Doğası Üzerine Eğilim Gösteren Temalar: VOSviewer ile Bibliyometrik Bir Analiz

**Özet:**

Bu araştırmada, fen eğitiminde bilimin doğası üzerine eğilim gösteren temaların belirlenmesi amaçlanmıştır. Bu amaç doğrultusunda Web of Science-WoS veri tabanı kullanılmıştır. Bilimsel inovatif alanındaki araştırma eğilimlerini tanımlamak ve sistematik bir şekilde haritalamak ve bu ağdaki entelektüel ilişkileri belirlemek amacıyla bibliyometrik yaklaşım benimsenmiş ve VOSviewer yazılım programı kullanılmıştır. Bu kapsamda çalışmaya 2013-2023 yılları baz alınmış, WoS veri tabanında Education & Educational Research' ve 'Education Scientific Disciplines' kategorileri seçilmiş ve sonuçların dökümünde elde edilen diğer yayın kategorileri çıkarılarak sadece makale türündeki çalışmalar veri setine dahil edilerek nihayetinde toplamda 263 makale incelenmiştir. Elde edilen bulgulara göre, 'Education & Educational Research' WoS kategorisinin ön planda olduğu, yıllara göre yapılan yayınların artıp azalan bir eğilim gösterdiği, en fazla çalışma yapan yazar Erduran S., en fazla atıf alan yazarın Lederman, N. G ve aralarında toplam bağlantı gücü en fazla yazarın ise Erduran, S olduğu, en fazla yayının 'Research in Scıence Educatıon' dergisinde yer aldığı, anahtar kelimelerden "nature of science", "science education" den sonra en sık "scientific literacy" ve "history of science" kelimelerinin; makalelerin özetlerinde yer alan en sık tekrar eden kelimelerin ise questionnaire ve interview kelimelerinin kullanıldığı ve son olarak ise toplam bağlantı gücü açısından ilk üçte yer alan ve aynı zamanda en fazla atıf alan ülkeler sırasıyla ABD, İngiltere ve Türkiye olarak tespit edilmiştir.

*Anahtar kelimeler*: Bilimin doğası, fen eğitimi, bibliyometrik analiz, VOSviewer, bilim tarihi ve bilimsel okuryazarlık.

## References

Abd-El-Khalick, F., & Lederman, N. G. (2000). Improving science teachers' conceptions of nature of science: A critical review of the literature. *International Journal of Science Education*, *22*(7), 665-701. https://doi.org/10.1080/095006900050044044

Adsiz, E., & Yiğit Kutluca, A. (2023). Teachers' views on the teaching of science process skills through integrated science and mathematics activities: Effect of the nature of science understandings. *Educational Academic Research, 48*(1), 27-41. https://doi.org/10.5152/AUJKKEF.2023.1023067

Agrifoglio, R., Metallo, C., & di Nauta, P. (2021). Understanding knowledge management in public organizations through the organizational knowing perspective: a systematic literature review and bibliometric Analysis. *Public Organization Review, 21*, 137–156. https://doi.org/10.1007/s11115-020-00480-7

American Association for the Advancement of Science (AAAS). (1993). Benchmarks for science literacy: A project 2061 report. Oxford University.

Aria, M., & Cuccurullo, C. (2017). Bibliometrix: an R-tool for comprehensive science mapping analysis. *Journal of Informetrics, 11*(4), 959-975. http://dx.doi.org/10.1016/j.joi.2017.08.007

Ayvacı, H. Ş., & Akdemir, E. (2017). An examination of theses in the field of nature of science published from 2013 in terms of different variables. *YYU Journal of Education Faculty*, *14*(1), 1178-1218. http://dx.doi.org/10.23891/efdyyu.2017.43

Bell, R. L., & Clair, T. L. S. (2015). Too little, too late: Addressing the nature of science in early childhood education. In K. C. Trundle & M. Saçkes (Eds.). *Research in early childhood science education* (pp. 125–141). Springer. https://doi.org/10.1007/978-94-017-9505-0_6

Bett, N. N., Piccolo, C., Roberson, N. D., Charbonneau, A. J., & Addison, C. J. (2023). Students' views on the nature of science in an interdisciplinary first-year science program: Content analysis of a weekly reflection activity. *Teaching & Learning Inquiry*, *11*. https://doi.org/10.20343/teachlearninqu.11.10

Bugingo, J. B., Yadav, L. L., Mugisha, I. S., & Mashood, K. K. (2024). Improving teachers' and students' views on nature of science through active instructional approaches: a review of the literature. *Science & Education, 33*, 29–71. https://doi.org/10.1007/s11191-022-00382-8

Caramaschi, M., Cullinane, A., Levrini, O., & Erduran, S. (2022). Mapping the nature of science in the Italian physics curriculum: From missing links to opportunities for reform. *International Journal of Science Education, 44*(1), 115–135. https://doi.org/10.1080/09500693.2021.2017061

Cheung, K. K. C., & Erduran, S. (2023). A systematic review of research on family resemblance approach to nature of science in science education. *Science & Education, 32*, 1637–1673. https://doi.org/10.1007/s11191-022-00379-3

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

CoHE. (2018). Fen bilgisi öğretmenliği lisans programı. https://www.yok.gov.tr/Documents/Kurumsal/egitim_ogretim_dairesi/Yeni-Ogretmen-Yetistirme-Lisans-Programlari/Fen_Bilgisi_Ogretmenligi_Lisans_Programi.pdf

Crowther, D. T., Lederman N.G. & Lederman, J. S. (2005). Methods and strategies: understanding the true meaning of nature of science. *Science and Children, 43*(2), 50-52. https://my.nsta.org/resource/7842/methods-and-strategies-understanding-the-true-meaning-of-nature-of-science

Çakıcı, Y. (2009). A prerequisite in science education: Understanding nature of science. *M.Ü. Atatürk Eğitim Fakültesi Eğitim Bilimleri Dergisi, 29,* 57-74. https://dergipark.org.tr/tr/pub/maruaebd/issue/370/2114

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

Effendi, D. N., Irwandani, Anggraini, W., Jatmiko, A., Rahmayanti, H., Ichsan, I. Z., & Rahman, M. M. (2021). Bibliometric analysis of scientific literacy using VOS viewer: Analysis of science education. *Journal of Physics: Conference Series*, *1796*, Article 012096. https://doi:10.1088/1742-6596/1796/1/012096

Gören, D. & Kaya, E. (2023). The relationship between students' understanding of nature of science and their metacognitive awareness. *Boğaziçi University Journal of Education. 40-2*(2). 315-336. https://doi.org/10.52597/buje.1123164

Gülmez Güngörmez, H. & Akgün, A. (2020). Student's opinions about nature of science activities embedded in process oriented guided inquiry learning method. *Electronic*

*Journal of Education Sciences, 9*(18), 128-148. https://dergipark.org.tr/en/pub/ejedus/issue/58123/718234

Hansson, L., & Leden, L. & Thulin, S. (2020). Book talks as an approach to nature of science teaching in early childhood education. *International Journal of Science Education, 42*(12), 2095-2111. https://doi.org/10.1080/09500693.2020.1812011

Hwang, G. J., Lai, C. L., & Wang, S. Y. (2015). Seamless flipped learning: A mobile technology-enhanced flipped classroom with effective learning strategies. *Journal of Computers in Education, 2*, 449-473. https://doi.org/10.1007/s40692-015-0043-0

Jaenudin, I., Aisyiah, P.N., Suryani, R., & Widodod, A. (2021). Content analysis of the nature of science on elementary thematic textbooks 2013 curriculum. *Premiere Educandum: Jurnal Pendidikan Dasar dan Pembelajaran, 11*(2), 227 – 242. https://doi.org/10.25273/pe.v11i2.9187

Jimenez-Aleixandre, M. P., Rodriguez, A. B., & Duschl, R. A. (2000). “Doing the lesson” or “doing science”: Argument in high school genetics. *Science Education, 84*(6), 757–792. https://doi.org/10.1002/1098-237X(200011)84:6<757::AID-SCE5>3.0.CO;2-F

Karataş Öztürk, S., Işık, H. & Alpaslan, M. M. (2023). Examination of middle school students’ views on the nature of scientific knowledge in the context of learning activities. *The Journal of Buca Faculty of Education, 57*,1555-1580. https://doi.org/10.53444/deubefd.1274069

Kaya, E. & Erduran, S. (2016). Yeniden kavramsallaştırılmış “Aile Benzerliği Yaklaşımı”: Fen eğitiminde bilimin doğasına bütünsel bir bakış açısı. *Part B: Türk Fen Eğitimi Dergisi*, *13*(2), 77-90.

Khishfe, R., & Abd-El-Khalick, F. (2002). Influence of explicit and reflective versus implicit inquiry-oriented instruction on sixth graders’ views of nature of science. *Journal of Research in Science Teaching, 39*(7), 551–578. http://doi.org/10.1002/tea.10036

Kurt, G., & Kaya, E. (2023). Reconceptualized family resemblance to the nature of science in the science curriculum and teachers’ views. *The Journal of Buca Faculty of Education, 57*, 1780-1807. https://doi.org/10.53444/deubefd.1289427

Kurtuluş, M. A., & Bilen, K. (2021). A bibliometric analysis on nature of science: a review of the research between 1986-2019. *Scientific Educational Studies, 5*(1), 47-65. https://doi.org/10.31798/ses.941238

Lederman, N. G. (2007). Pedagogy and the practice of science. *International Journal of Science Education*, *29*(7), 931-934. https://doi.org/10.1080/14697680601043290

Li, B., & Xu, Z. (2022). A comprehensive bibliometric analysis of financial innovation. *Economic Research-Ekonomska Istraˇ zivanja, 35*(1), 367–390. https://doi.org/10.1080/1331677X.2021.1893203

Librea-Carden, M. R., Mulvey, B. K., Borgerding, L. A., Wiley, A. L., & Ferdous, T. (2021). 'Science is accessible for everyone': Preservice special education teachers' nature of science perceptions and instructional practices. *International Journal of Science Education*, *43*(6), 949–968. https://doi.org/10.1080/09500693.2021.1893857

Librea-Carden, M.R., & Mulvey, B.K. (2023). The potential of nature of science (NOS) in special education (SPED): Preservice teachers' conceptions, plans, and identified NOS implications for SPED. *Research in Science Education*, 53, 1097–1118. https://doi.org/10.1007/s11165-023-10125-6

Maryanti, R., Nandiyanto, A. B. D., Hufad, A., Sunardi, S. Husaeni, D. N. A., & Husaeni, D. F. A. (2023). A computational bibliometric analysis of science education research using vosviewer. *Journal of Engineering Science and Technology*, *18*(1), 301- 309. https://jestec.taylors.edu.my/Vol%2018%20Issue%201%20February%20%202023/18_1_20.pdf

McComas, W. F. (2004). Keys to teaching the nature of science: Focusing on the nature of science in the science classroom. *The Science Teacher*, *71*(9), 24-27. https://www.researchgate.net/publication/262450010

NGSS Lead States. (2013). *The next generation science standards*. National Academy Press.

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. *Annals of Internal Medicine*, *151*(4), 264–269. https:// doi. org/10. 7326/ 0003- 4819-151-4- 20090 8180- 00135

MoNE. (2018). *Fen bilimleri dersi öğretim programı (ilkokul ve ortaokul 3, 4, 5, 6, 7 ve 8. sınıflar)*.

NRC (National Research Council). (1996). National science education standards. *National Academy of Sciences*. https://doi.org/10.17226/4962

Okan, B., & Kaya, E. (2023). A content analysis of the representation of the nature of science in a Turkish science textbook. In: Carvalho, G.S., Afonso, A.S., Anastácio, Z. (eds) *Fostering scientific Citizenship in an Uncertain World. Contributions from Science Education Research*, 13. Springer. https://doi.org/10.1007/978-3-031-32225-9_5

Ozan, F., & Uluçınar Sağır, Ş. (2020). The effect of stem activities to the secondary school students opinions on the nature of science. *Journal of STEAM Education, 3* (2), 32-43. https://dergipark.org.tr/en/download/article-file/1188146#764293]-1188146.pdf

Öztürk, E., & Kaptan, F. (2014 A view on "ESERA 2009" science education researches conferance and its content studies about nature of science, argumentation, history, philosophy and sociology of science. *Kastamonu Education Journal, 22*(2), 649-672. https://dergipark.org.tr/tr/pub/kefdergi/issue/22602/241502

Pradana, M., Elisa, H. P., & Syarifuddin, S. (2023). The growing trend of Islamic fashion: A bibliometric analysis. *Cogent Social Sciences*, *9*(1), 2184557. https://doi.org/10.1080/23311886.2023.2184557

Prahani, B. K**.,** Rizki, I. A., Suprapto, N., Irwanto, I., & Kurtulus, M. A. (2024). Mapping research on scientific creativity: A bibliometric review of the literature in the last 20 years. *Thinking Skills and Creativity, 52*, 101-495. https://doi.org/10.1016/j.tsc.2024.101495

Pranckute, R. (2021). Web of Science (WoS) and scopus: The titans of bibliographic information in today's academic world. *Publications*, 9, 12. https://doi.org/10.3390/publications9010012

Sarkar, A., Wang, H., Rahman, A., Memon, W. H., & Qian, L. (2022). A bibliometric analysis of sustainable agriculture: based on the Web of Science (WOS) platform. *Environmental Science and Pollution Research, 29,* 38928–38949. https://doi.org/10.1007/s11356-022-19632-x

Solehuddin, M., Muktiarni, M., Rahayu, N. I., & Maryanti, R. (2022). Counseling guidance in science education: Definition, literature review, and bibliometric analysis. *Journal of Engineering Science and Technology, (Special Issue on ISCoE2022),* 1- 13. https://jestec.taylors.edu.my/Special%20Issue%20ISCoE%202022_1/ISCoE%201_01.pdf

Sönmez, H., & Hastürk, G. H. (2020). Bibliographic analysis of doctoral level thesis studies in science education domain in Turkey. *Journal of Human Science Researches , 9*(5), 3174-3194. https://dergipark.org.tr/tr/download/article-file/1099718

Stadermann, H. K. E., & Goedhart, M. J. (2020). Secondary school students' views of nature of science in quantum physics. *International Journal of Science Education, 42*(6), 997–1016. https://doi.org/10.1080/09500693.2020.1745926

Suryani, R., Widodo, A., & Sujana A. (2022). Content analysis of the nature of science in the thematic learning book for lower grade elementary school 2013 curriculum. *Proceeding The 4th International Conference on Elementary Education, 4*(1), 596-608.

Takriti, R. A., Erduran, S., Tairab, H., Kaya, E., Alhosani, N., Rabbani, L., & Alamirah, I. (2024). Translation and validation of a questionnaire for measuring teachers' views on nature of science. *Research in Science &Technological Education*, *42*(3), 812–827. https://doi.org/10.1080/02635143.2022.2138846

Taşkın, T. (2021). An examination on nature of science studies in terms of different variables. *Bolu Abant İzzet Baysal University Journal of Faculty of Education, 21*(1), 1-20. https://dx.doi.org/10.17240/aibuefd.2021.21.60703-794024

Tosun, C. (2024). Analysis of the last 40 years of science education research via bibliometric methods. *Science & Education, 33,* 451–480. https://doi.org/10.1007/s11191-022-00400-9

Wahbeh, N., & Abd-El-Khalick, F. (2014). Revisiting the translation of nature of science understandings into instructional practice: Teachers' nature of science pedagogical content knowledge. *International Journal of Science Education*, *36*(3), 425–466. https://doi.org/10.1080/09500693.2013.786852

Widowati, A., Widodo, E., Anjarsari, P., & Setuju (2017). The development of scientific literacy through nature of science (NoS) within inquiry based learning approach. *Journal of Physics: Conference Series*, *909*, Article 012067. https://doi.org/10.1088/1742-6596/909/1/012067

Van Eck, N. J., & Waltman, L. (2022). VOSviewer manual (version 1.6.18. CWTS Meaningful Metrics. https://www.vosviewer.com/documentation/Manual_VOSviewer_1.6.8.pdf

Voss, S., Kent-Schneider, I., Kruse, J., & Daemicke, R. (2023). Investigating the development of preservice science teachers' nature of science instructional views across rings of the family resemblance approach wheel. *Science & Education 32*,1363–1399 https://doi.org/10.1007/s11191-023-00418-7

Yacoubian, H. A. (2021). Students' views of nature of science. *Science & Education, 30*, 381–408. https://doi.org/10.1007/s11191-020-00179-7

Yanuarti, E.A., & Suprapto, N. (2021). Ten years of research on history of science (physics): A bibliometric analysis. *Studies in Philosophy of Science and Education*, *2*(1), 7-16. https://doi.org/10.46627/sipose.v2i1.66

Necatibey Faculty of Education Electronic Journal of Science and Mathematics Education
Vol. 18, Issue 2, December 2024, pp. 503-533. ISSN: 1307-6086

Research Article

# A Scale Development Study to Use the Flipped Classroom Model in Mathematics Education*

**Emine ÖZDEMİR[1], Beyzanur TOPÇU GÖRE[2]**

[1] Balıkesir University, Necatibey Faculty of Education, Türkiye, eozdemir@balikesir.edu.tr, http://orcid.org/0000-0002-4114-0005

[2] Balıkesir Karesi District Directorate of National Education, Türkiye, beyzanet26@gmail.com, http://orcid.org/0000-0003-2826-5601

Received : 29.05.2024 Accepted : 27.11.2024

Doi: https://doi.org/10.17522/balikesirnef.1488745

---

*Abstract* – In this study, it was aimed to develop a valid and reliable scale to evaluate the use of the Flipped Classroom Model in secondary school mathematics teaching. The study design was determined as a survey design. In the process of developing the scale, an item pool was first created and expert opinion was obtained. Adjustments were made according to the opinions of experts. The study group was determined by simple random sampling and the application was carried out with a total of 871 middle school students from grades 5, 6, 7, and 8. According to the data obtained, validity and reliability analyses of the scale were conducted with SPSS and AMOS programs. The Flipped Classroom Model Evaluation (FCME) Scale consists of two subscales: the "Evaluation of Teaching Conducted with Video Lessons" (ETCVL) subscale and the "Evaluation of Teaching Conducted with Face-to-Face Lessons" (ETCFFL) subscale. As a result of the analyses, the FCME scale consisting of 32 items in total was developed. The reliability coefficient of the ETCVL subscale was .846, the reliability coefficient of the ETCFFL subscale was .775 and the reliability coefficient of the scale as a whole was .886. As a result of the analysis, it was determined that the FCME Scale is valid and reliable.

*Keywords:* Flipped classroom model, math education, secondary school students, scale development.

------------

Corresponding author: Emine ÖZDEMİR , eozdemir@balikesir.edu.tr

---

*This study includes a part of the second author's Master's thesis entitled "A Scale Development Study to Use the Flipped Classroom Model" conducted under the supervision of the first author.

## Introduction

With the rapid development of science and technology, it is known that the use of smartphones, tablets, computers, etc. has increased and products such as videos, photographs, and designs are created using these tools and shared on various social platforms (Kocaman Karoğlu, 2015). In societies exposed to this situation, it is known that there is an increase in curiosity about technology, active use of technology, and an increase in the desire to improve their knowledge and skills on how to use technology. While technology is used so actively in daily life, the fact that educational environments lag in the use of technology has led to digital incompatibility (Atal & Koçak Usluel, 2011).

The quality of the education provided is directly proportional to the quality of the education provided by teachers blending technology, pedagogy, and field knowledge. In the literature, the blending of these three knowledge is called Technological Pedagogical Content Knowledge (TPACK) (Mutluoğlu & Erdoğan, 2016). When we look at the common intersection set of the components, TPACK defines the course designs prepared by integrating technological-pedagogical-field knowledge. (Koehler & Mishra, 2009). In the updated curricula by the Ministry of National Education (MoNE), the necessity of imparting eight key competencies to students during the educational processes is mentioned, including digital competence, competence in science and technology, learning to learn, and mathematical competence (MoNE, 2018).

The FC Model is a learning model that emerged to present the subject to the student with formats such as slides, videos, interactive content, etc. on digital platforms so that students who cannot come to class for various reasons do not fall behind (Baker, 2000). According to this learning model, students learn the subject through digital platform content and then come to the educational environment to reinforce the subject with activities. At first, this method was used only for students who could not come to class, but after receiving positive feedback, the researchers started to apply this method to the whole class. With this method, while saving the time allocated for lecturing in the lesson, it also created an opportunity for students who lacked knowledge or missed the subject to complete the subject (Bender, 2018; Bergmann, 2011).

When the studies in the literature are examined, it is seen that the lessons taught with the FC model increase students' academic achievement, active participation in the classroom, self-regulation skills, retention of the subject, teacher-student communication, make the lesson more understandable, fun and remarkable, and students begin to develop positive attitudes towards mathematics lessons (Arslan, 2021; Bolatlı, 2018; Bulut, 2019; Güç, 2017; Houston, 2020; Özdemir, 2016; Wei et al., 2020).

When the studies on the use of the FC Model were examined, it was found that there was no study in which students from all grade levels were studied at the secondary school level, and that interview forms were used as data collection tools to determine the opinions of students, and that there was no scale with proven validity and reliability. It is thought that the research conducted in this direction can contribute to these deficiencies identified in the literature.

It was determined that there were few studies on the use of the FC Model in secondary school mathematics courses and the problem situations (attitudes, opinions, academic achievement, etc.) examined in the studies were similar. Considering the study group, it was observed that the study was generally conducted with a single grade level. Upon examination, it was determined that there was no scale for the evaluation of the FC Model and that there was no study conducted with all grade levels at the same time. Accordingly, this study aims to develop a scale with proven validity and reliability for middle school students to evaluate the use of the FC model in mathematics education and to contribute to the deficiencies in the literature.

## Method

### Research Design

As the aim of this study is to develop a scale, the research design is a survey design from quantitative research methods. The survey design enables the quantitative presentation of the attitudes, opinions or tendencies of the population by conducting studies with a sample selected from the population (Creswell, 2017) and obtaining information by reaching many participants (Büyüköztürk et al., 2018).

### Participants

The schools and classes to be included in the study group were determined by a simple random sampling method. According to this sampling method, each individual is equally likely to be selected for sampling and it is the most powerful and valid method compared to other sampling methods in terms of representing the universe (Büyüköztürk et al., 2018).

The study group of the research included a total of 871 secondary school students from six public schools in Sancaktepe district of Istanbul province, studying in grades 5-6-7-8 in the second semester of the 2020-2021 academic year. Of these 871 students, 512 were female and 359 were male. In terms of grade level, there are 236 students in 5th grade, 289 students in 6th grade, 200 students in 7th grade, and 146 students in 8th grade.

## Data Collection

### *Process of Scale Development*

In the process of scale development in line with the purpose of the study, the process in Figure 1 was followed based on Yurgudül's (2005) scale development scheme.

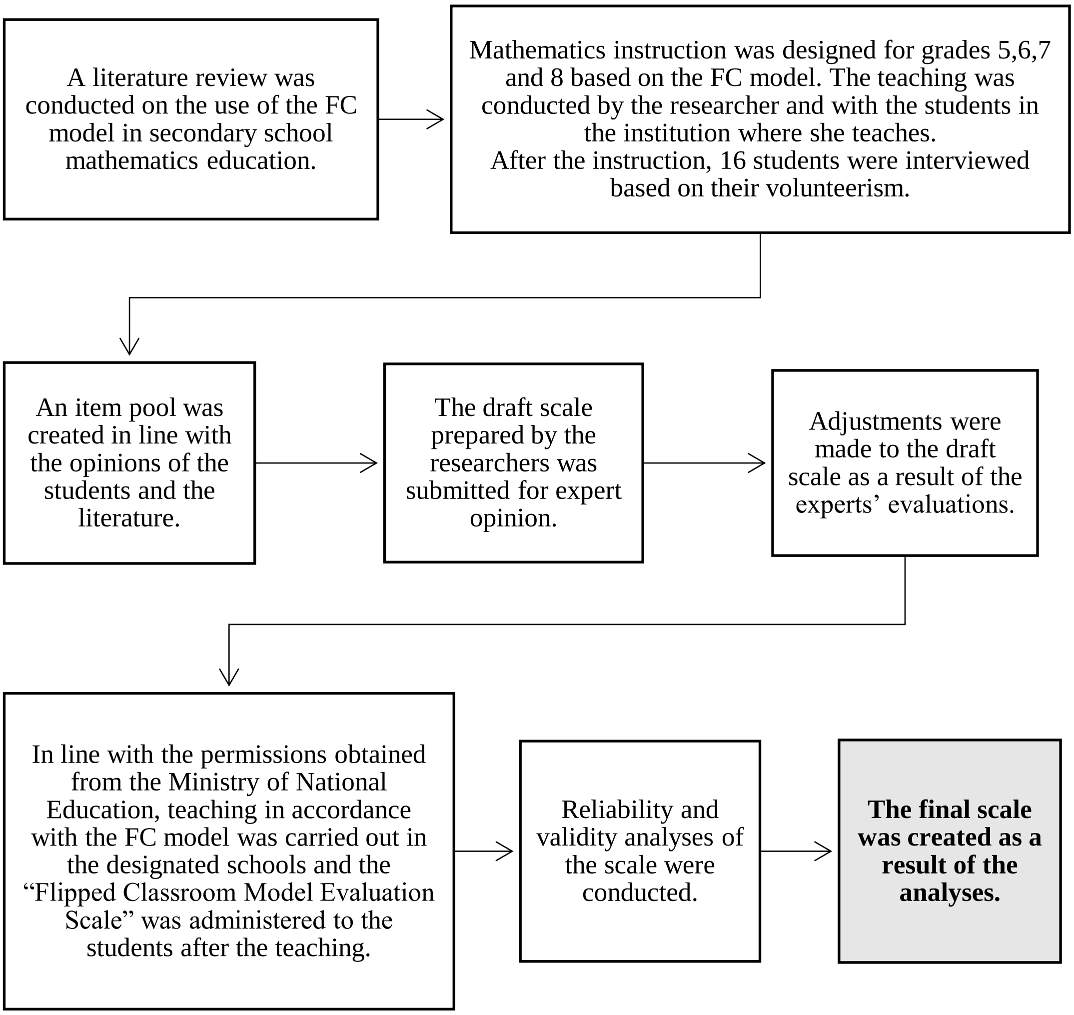



**Figure 1** Scale Development Process

In the national literature, graduate thesis studies on scale development were examined through the YÖKTEZ database. Since this study is a scale development study for the evaluation of a FC model and the AMOS program of structural equation modeling was used, the combinations of the keywords “flipped, scale development, structural equation, AMOS” were searched by keeping the keyword “scale development” constant.

The literature review was limited with the criteria of being between 2012 and 2022 to cover the last 10 years and being in the field of education and training. When the words “flipped, scale development, AMOS”, “flipped, scale development, structural equation” were searched in the thesis name, abstract, subject and index sections of the theses from the detailed search section provided by YÖKTEZ, no study was found. Then, when the words “flipped, scale development” were searched, three thesis studies in which a scale was developed related to the FC model were found and one article was found as a result of the literature review. When the words “structural equation, scale development” were searched, 14 thesis studies were found. Six of these studies were excluded from the literature review because they did not use structural equation modeling during scale development.

In addition to these studies, the study titled “Scale development process in the field of education: a content analysis study” by Şahin and Boztunç Öztürk (2018), which examined 69 articles developed in the field of education in Türkiye between 2010 and 2016, was also included in the review. For this reason, a total of 13 studies were analyzed under the title of “studies on scale development”.

In the literature review, four scales related to the flipped classroom model were found. One of these studies is Durak's (2017) adaptation of the scale developed by Hao (2016) to measure students' readiness into Turkish. In his study, Kurtoğlu (2019) used the scale adapted into Turkish by Durak (2017) and adapted it to teachers by conducting reliability studies. Erensayın (2019) developed self-efficacy and perception scales for teachers regarding the applicability of the FC model; Akgün (2015) developed a scale aiming to examine the effect of the FC model on students' academic achievement and opinions.

When the thesis studies in which the structural equation model was used among the scale development studies in the field of education were examined, it was determined that the analysis program used varied. It can be said that the use of different analysis programs such as AMOS and LISREL in the studies depends on the preference of the researchers.

In the thesis study conducted by Balkaya (2022), it was aimed to develop the “YouTube Usage Scale” for middle school students. For this purpose, a total of 644 students from 5th, 6th, 7th and 8th grades were included in the study. The draft scale, which was initially prepared with 47 items, was reduced to 42 items after expert opinion. Then, EFA and CFA analyses were conducted with SPSS and AMOS programs. As a result of these analyzes, the final scale was determined as 25 items. The reliability coefficient of the final scale was found to be .91.

In the study conducted by Nanto (2021), the effect of error management culture and job attraction behaviors on organizational creativity was examined. Accordingly, the researcher developed the "Error Management Culture" scale for quantitative data. The sample of the study consisted of 747 primary and secondary school teachers. For the scale, a 41-item draft scale was first created as a result of the literature review. The number of items was reduced to 26 by removing 15 items from the draft scale sent to expert opinion. Then, EFA and CFA analyses conducted with SPSS and AMOS programs resulted in a final scale consisting of 16 items. The reliability coefficient of the scale was found to be .846.

The validity and reliability analyses of the scales developed by Erensayın (2019), Kurtoğlu (2019) and Durak (2017) were included in the studies. However, in the scale study developed by Akgün (2015) for the evaluation of the FC model, it was found that no information on validity and reliability analyses was provided.

The draft scale was prepared in a 5-point Likert scale and named the "Flipped Classroom Model Evaluation (FCME) Scale". Due to the structure of the FC model, teaching takes place in two phases: "video lesson teaching" and "face-to-face teaching". Due to this structure of the FC model, the FCME scale consists of two subscales, namely the "Evaluation of Teaching Conducted with Video Lessons (ETCVL) Subscale" and the "Evaluation of Teaching Conducted with Face-to-Face Lessons (ETCFFL) Subscale". This nomenclature was chosen because the scale consists of two subscales.

***Process of Creating the Substance Pool***

In accordance with the decision taken at the meeting of Balıkesir University Science and Engineering Sciences Ethics Committee dated 20.04.2021 and numbered 2021/2, it was approved to collect the data of the study.

In the training given to create an item pool, content was prepared in Edpuzzle according to the gains below and teaching was carried out.

- M.5.3.1.2. Collects data related to research questions and shows them in a frequency table and column graph.

- M.6.1.5.3. Makes multiplication of a natural number and a fraction and makes sense of it.
- M.6.1.5.4. Makes multiplication of two fractions and makes sense of it.
- M.7.1.3.2 Multiplies and divides rational numbers.
- M.8.1.3.4. Performs multiplication and division operations with square roots.

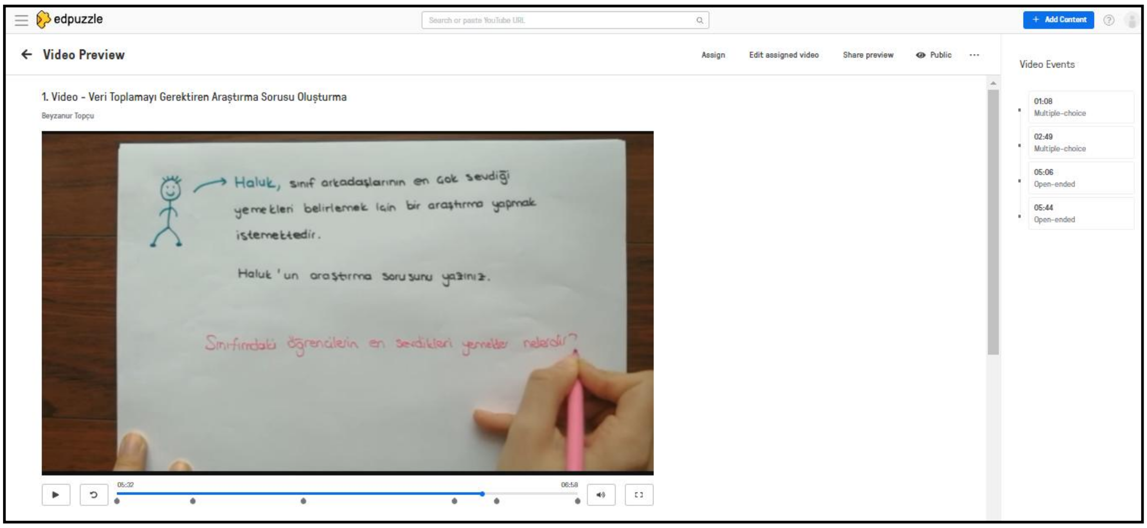



**Figure 2** An Example of Video Lessons

After the instruction, the step of creating an item pool for the scale was started. At this stage, volunteer students were contacted to get their opinions on the evaluation of teaching with the FC model. The following open-ended questions were asked to the 16 students who volunteered to give their opinions and the opinions of the students were obtained.

- “Has studying the topic beforehand had any effect on the desire to participate in solving problems during the lesson? If so, what kind of impact did it have?
- What are your thoughts on interactive questions in video lessons?
- What are your thoughts on feedback provided for interactive questions in video lessons?
- What do you think about using Edpuzzle and online quizzes/games at the end of video lessons? What are the positive and negative aspects?
- How did you feel using a site like Edpuzzle, Wordwall, Quizizz, etc., which you had never used before?

- Can you make a comparison between math lessons taught with the FC model and those not taught with it?"

*"In this application, I think it is better because there is not a specific time like a lecture because some of us may be in another place, for example, they may not have internet, so they cannot open it, so it is good that there is not a specific time, they watch it the next day and then they go to class (Student C)"*

Based on this student's opinion, items 1, 2, and 21 of the ETCVL subscale were created.

*"When someone gets it right, you write the correct answer bravo, and when they get it wrong, they can see why they got it wrong. I think this is very nice (Student N)."*

Similarly, Student N's opinions helped to write the 4th and 8th items of the ETCVL Subscale.

When examining the qualitative data resulting from the interviews, it was observed that students responded to the "evaluation of teaching conducted with video lessons", "evaluation of teaching conducted face-to-face", "the use of technology in mathematics instruction", and "their attitudes towards mathematics" within the FC model. When the responses from the students were analyzed, the idea that the evaluation of teaching should consist of two stages emerged.

These two stages are named "Evaluation of Teaching Conducted with Video Lessons" and "Evaluation of Teaching Conducted Face-to-Face," and accordingly, an item pool has been created. At the stage of creating the item pool, the scale was prepared in 5-point Likert type from multiple scales. In the Likert scale, various statements are directed to the sample and it is expected to determine the degree of agreement or disagreement with these statements (Altunışık et al., 2005). After reviewing the responses, the draft scale was prepared as 90 items. These items were prepared and organized using the Microsoft Form tool.

***Content Validity***

It was aimed to ensure the content validity of the scale by applying expert opinion. A form was prepared in which "Appropriate, Should be improved, Not appropriate" options were added to each of the scale items. If the option "should be improved" was selected, a text box was opened and the experts were asked to indicate their suggestions.

The draft scale was sent via e-mail to a total of 11 experts, including 3 mathematics and 1 Turkish teacher with at least 5 years of experience, 1 measurement and evaluation, 1 Turkish education, 3 mathematics education, and 2 educational experts from other fields.

The content validity index of the draft scale was calculated based on the opinions of the experts. The Content Validity Ratio (CVR) approach is used in scale development studies to quantify the opinions of experts regarding the content validity of the items in the developed scale (Yurdugül & Bayrak, 2012). For each item in the draft scale, a calculation was made using the content validity ratio formula.

Since the number of experts was 11, the content validity criterion for each item was determined as .59 (Yurdugül, 2005). The CVR value was found to be higher than .59 for each item on the scale. With the content validity index (CVI), a value is calculated for the entire scale. CVI is calculated by averaging the CVR values of the items decided to be included in the scale (Yurdugül & Bayrak, 2012). The experts evaluated the 90 items from different perspectives, such as the FC model, grammar, suitability for the purpose, comprehensibility of the items, and the presence of similar items. Edits were made by taking into account the opinions of the experts and 26 of the 90 items were deleted and a draft scale consisting of 64 items was formed. Accordingly, at $\alpha = .05$ level of significance, the content validity index value of the FCME scale was calculated as 0.98 and it was concluded that the content validity of the scale had a high value.

Scoring for the scale items was done as 1=Strongly disagree, 2=Disagree, 3=Neither agree nor disagree, 4=Agree, 5=Strongly agree for the positive items, and the reverse for the negative items.

Students who scored high on the scale can be said to have positive evaluations of mathematics teaching according to the FC model.

***Teaching Process Based on the FC Model and Collection of Scale Data***

In the six public secondary schools included in the study, mathematics teaching based on the FC model was implemented. Before the implementation, care was taken to ensure that there were students from all levels (5th, 6th, 7th, and 8th grades) of each secondary school included in the implementation. Instruction was carried out with a total of 871 students and then the instruction was evaluated by applying the draft scale.

In the process of teaching based on the FC model, content was prepared according to the following acquisitions and teaching was carried out.

- M.5.2.2.3.Determines and draws the elements of rectangle, parallelogram, rhombus and trapezoid.
- M.6.3.2.2.Form the area relation of a parallelogram and solve related problems.
- M.7.3.2.1.Explain the side and angle properties of regular polygons.
- M.7.3.2.2.Determines the diagonals, interior and exterior angles of polygons, calculates the sum of the measures of interior and exterior angles.
- M.8.3.1.3.Relates the side lengths of a triangle to the measures of the angles opposite these sides.

**Data Analysis**

Structural Equation Modeling was preferred in the data analysis of the study. Structural Equation Modeling (SEM) adopts a confirmatory approach by analyzing all direct or indirect effects between variables at the same time, while also taking into account the errors of the variables (Demir, 2016).

After the implementation of mathematics teaching according to the FC model, the draft scale was applied to the students. The data received after the application were transferred to the computer environment. Data analysis was conducted using the SPSS 24.0 program for factor analysis, t-test, and reliability analysis, and the AMOS 24.0 program for factor analysis within the scope of SEM.

Validity analyses (exploratory factor analysis, confirmatory factor analysis) and reliability analyses (item analysis, t-test, Cronbach's alpha coefficient) were conducted separately for the two subscales of the FCME scale, namely, "Evaluation of Teaching Conducted with Video Lessons (ETCVL) subscale" and "Evaluation of Teaching Conducted with Face-to-Face Lessons (ETCFFL) Subscale".

## Results

***Analyses Related to the Evaluation of Teaching Conducted with Video Lessons (ETCVL) Subscale***

The ETCVL subscale was determined as 49 items after expert opinion. First, the Kaiser Meyer Olkin (KMO) value was calculated to determine the suitability of this subscale for factor analysis. Since the KMO value of the ETCVL subscale was found to be .93, it was determined that the scale data were suitable for factor analysis.

Five exploratory factor analyses were applied to the ETCVL subscale and a total of 21 items with loadings below .50 or overlapping items were removed from the scale.

**Table 1** Results of the Rotated Components Matrix of the Items in the ETCVL Subscale

| | Rotated component matrix | | | | | |
|---|---|---|---|---|---|---|
| | Factors | | | | | |
| Items | 1 | 2 | 3 | 4 | 5 | 6 |
| videolesson14 | .720 | .021 | -.054 | .106 | .076 | -.016 |
| videolesson13 | .708 | .028 | .054 | .046 | .033 | -.054 |
| videolesson12 | .706 | .031 | .156 | .069 | -.010 | -.037 |
| videolesson23 | .682 | .060 | .011 | .083 | -.039 | -.122 |
| videolesson26 | .676 | .112 | .291 | .122 | -.003 | -.014 |
| videolesson25 | .659 | .191 | .191 | .168 | .057 | -.029 |
| videolesson22 | .652 | .083 | .215 | .127 | .018 | -.030 |
| videolesson16 | .626 | -.119 | -.007 | .209 | .149 | .061 |
| videolesson11 | .612 | .016 | .152 | .178 | .085 | .021 |
| videolesson9 | .579 | .005 | .141 | .092 | .267 | .096 |
| videolesson10 | .573 | .059 | .188 | .092 | .183 | .071 |
| videolesson4 | .558 | .083 | .090 | .039 | .252 | -.101 |
| videolesson18 | .262 | .723 | -.196 | -.022 | -.216 | .007 |
| videolesson30 | .096 | .713 | -.010 | .057 | -.037 | .120 |
| videolesson19 | .223 | .643 | -.019 | .059 | -.174 | .114 |
| videolesson35 | .153 | .643 | .136 | -.125 | -.050 | .107 |
| videolesson29 | -.155 | .610 | .222 | -.067 | .188 | .077 |
| videolesson31 | -.279 | .594 | .138 | .027 | .233 | .040 |
| videolesson34 | .253 | .065 | .792 | .096 | .026 | -.017 |
| videolesson37 | .193 | .081 | .781 | .161 | -.031 | -.033 |
| videolesson33 | .407 | .055 | .587 | .084 | .048 | .029 |
| videolesson40 | .176 | -.069 | .127 | .750 | .022 | .066 |
| videolesson38 | .224 | .032 | .221 | .685 | .070 | -.147 |
| videolesson39 | .258 | -.011 | .000 | .668 | .162 | -.002 |
| videolesson2 | .213 | -.022 | .071 | .070 | .746 | -.053 |
| videolesson3 | .265 | -.085 | -.082 | .162 | .677 | -.010 |
| videolesson45 | -.009 | .198 | -.056 | -.022 | .067 | .799 |
| videolesson47 | -.090 | .179 | .032 | -.024 | -.127 | .787 |

It was determined that the lowest loading values of the items constituting the scale were .558 and the highest .799.

The items belonging to the ETCVL subscale were coded as “videolesson...”. In addition, it was found that the subscale of the ETCVL consisted of six factors. These factors were named “positive attitude” (items 4, 9, 10, 11, 12, 13, 14, 16, 22, 23, 25, 26), “negative attitude” (items 18, 19, 29, 30, 31, 35), “self-efficacy” (items 33, 34 and 37), "advantages" (items 38, 39 and 40), "accessibility" (items 2 and 3) and "safety" (items 45 and 47).

Confirmatory factor analysis was conducted to determine the construct validity of the remaining 28 items in the ETCVL subscale. As a result of the analysis, items 4, 29 and 31 were removed from the scale because their standardized regression weights were below .5. Improvement was made by creating covariance between items with high error percentages. Table 2 shows the fit values obtained as a result of these procedures.

**Table 2** Confirmatory Factor Analysis Fit Values of the Items in the ETCVL Subscale

| Model fit indices | | Good fit | Acceptable fit | Found value |
|---|---|---|---|---|
| CMIN/DF | | $x^2$/ sd ≤ 3 | $x^2$/ sd ≤ 5 | 2.520 |
| Comparative fit indices | NFI | .95 ≤ NFI | .90 ≤ NFI | .906 |
| | CFI | .97 ≤ CFI | .95 ≤ CFI | .941 |
| | RMSEA | RMSEA ≤ .05 | RMSEA ≤ .08 | .042 |
| Absolute fit indices | GFI | .90 ≤ GFI | .85 ≤ GFI | .942 |
| Residual-Based Fit Indices | RMR | 0 < RMR ≤ .05 | 0 < RMR ≤ .08 | .041 |

According to the fit values of $x^2$/df = 2.520 < 3, .90 ≤ NFI = .906, .95 ≤ CFI = .941, RMSEA = .042 ≤ .05, .90 ≤ GFI = .942, RMR = .041 ≤ .05 for the final scale items of the ETCVL subscale, there is an excellent fit between the model and the data (Karagöz, 2021).

Based on the findings obtained as a result of the analyses, the validity of the final six-factor version of the ETCVL subscale was proved.

The AMOS diagram of the ETCVL subscale obtained as a result of confirmatory factor analysis is shown in Figure 2.

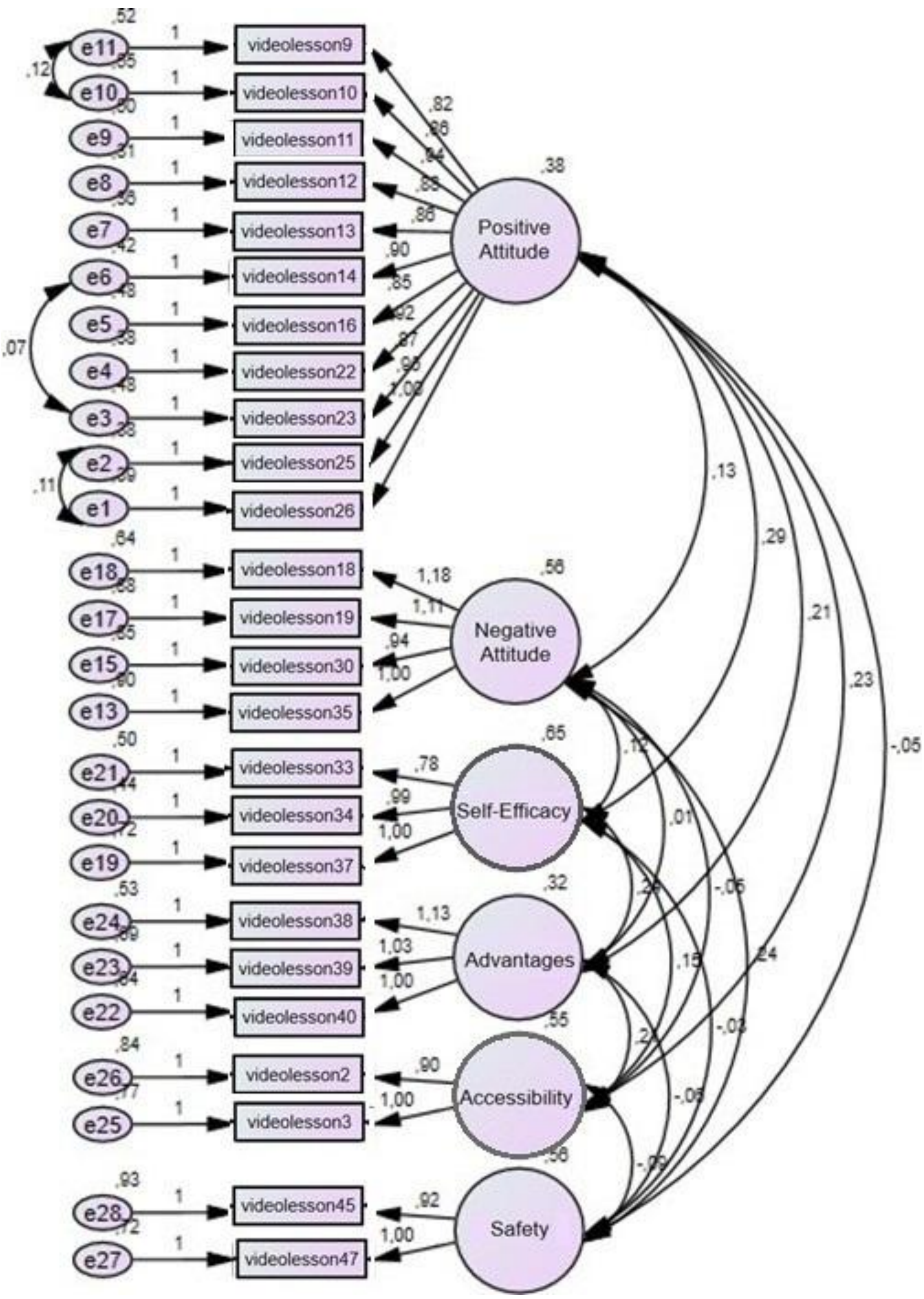



**Figure 2** AMOS Diagram of the Items in the ETCVL Subscale

### *Analyses Related to the Evaluation of Teaching Conducted with Face-to-Face Lessons (ETCFFL) Subscale*

The ETCFFL subscale was determined as 15 items after expert opinion. First, the Kaiser Meyer Olkin (KMO) value was calculated to determine the suitability of this subscale for factor analysis. Since the KMO value of the ETCFFL subscale was found to be .851, it was determined that the scale data were suitable for factor analysis.

Three exploratory factor analyses were applied to the ETCFFL subscale and a total of 6 items with loadings below .45 or overlapping items were removed from the scale.

**Table 3** Results of the Rotated Components Matrix of the Items in the ETCFFL Subscale

| Items | Rotated component matrix<br>Factors<br>1 | <br><br>2 |
|---|---|---|
| facetofacelesson7 | .781 | .231 |
| facetofacelesson6 | .759 | .168 |
| facetofacelesson5 | .722 | .229 |
| facetofacelesson10 | .704 | .057 |
| facetofacelesson13 | .600 | -.168 |
| facetofacelesson1 | .548 | .394 |
| facetofacelesson11 | .545 | -.109 |
| facetofacelesson2 | .025 | .891 |
| facetofacelesson3 | .058 | .882 |

It was determined that the lowest loading values of the items constituting the scale were .545 and the highest .891.

The items belonging to the ETCFFL subscale were coded as "facetofacelesson...". In addition, it was found that the subscale of the ETCFFL consisted of two factors. These factors were named "positive attitude" (items 1, 5, 6, 7, 10, 11, and 13) and "negative attitude" (items 2 and 3).

Confirmatory factor analysis was conducted to determine the construct validity of the remaining 9 items in the ETCFFL subscale. As a result of the analysis, items 11 and 13 were removed from the scale because their standardized regression weights were below .5. Improvement was made by creating covariance between items with high error percentages. Table 2 shows the fit values obtained as a result of these procedures.

**Table 4** Confirmatory Factor Analysis Fit Values of the Items in the ETCFFL Subscale

| Model fit indices | | Good fit | Acceptable fit | Found value |
|---|---|---|---|---|
| CMIN/DF | | $x^2$/ sd ≤ 3 | $x^2$/ sd ≤ 5 | 4.117 |
| | NFI | .95 ≤ NFI | .90 ≤ NFI | .976 |
| Comparative fit indices | CFI | .97 ≤ CFI | .95 ≤ CFI | .982 |
| | RMSEA | RMSEA ≤ .05 | RMSEA ≤ .08 | .060 |
| Absolute fit indices | GFI | .90 ≤ GFI | .85 ≤ GFI | .984 |
| Residual-based fit indices | RMR | 0 < RMR ≤ .05 | 0 < RMR ≤ .08 | .039 |

According to the fit values of $x^2$/df = 4.117 < 5, .95 ≤ NFI = .976, .95 ≤ CFI = .982, .05 ≤ RMSEA = .060, .90 ≤ GFI = .984, RMR = .039 ≤ .05 for the final scale items of the ETCFFL subscale, there is an excellent fit between the model and the data (Karagöz, 2020).

Based on the findings obtained as a result of the analyses, the validity of the final two-factor version of the ETCFFL subscale was proved.

The AMOS diagram of the ETCFFL subscale obtained as a result of confirmatory factor analysis is shown in Figure 3.

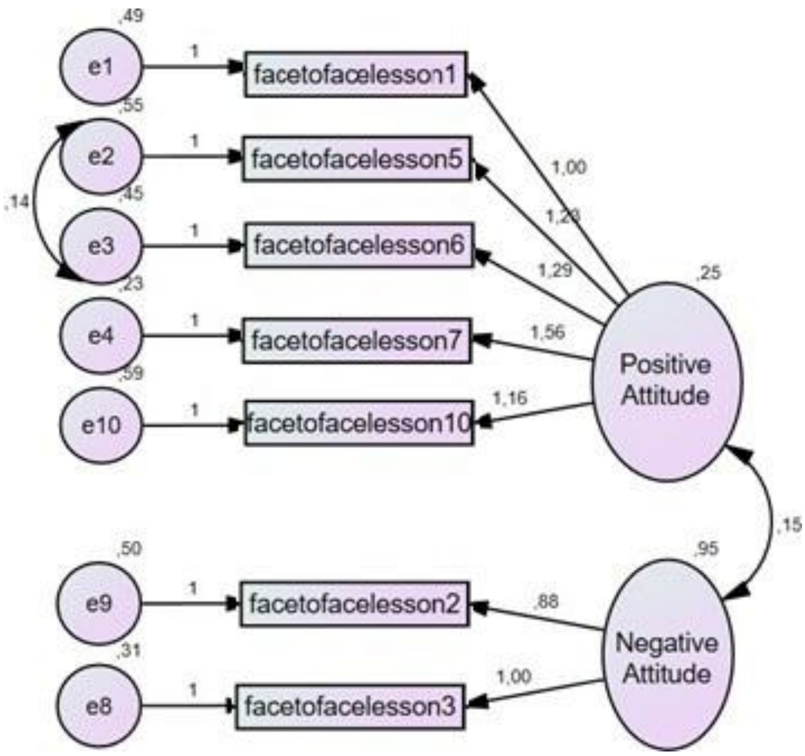



**Figure 3** AMOS Diagram of the Items in the ETCFFL Subscale

As a result of the exploratory and confirmatory factor analyses of the FCME scale,

- Twenty-four of the 49 items in the ETCVL subscale were removed from the scale, leaving 25 items in the final scale. Items 2, 3, 9, 10, 11, 12, 13, 14, 16, 18, 19, 22, 23, 25, 26, 30, 33, 34, 35, 37, 38, 39, 40, 45, 47 remained in the final version of ETCVL subscale.
- Eight of the 15 items in the ETCFFL subscale were removed from the scale and 7 items remained in the final scale. Items 1, 2, 3, 5, 6, 7 and 10 remained in the final version of the ETCFFL subscale

Table 5 and Table 6 show the factors formed in the two subscales of the final version of the FCME scale and the items included in these factors.

**Table 5** The Factors and Items in the ETCVL Subscale and the Items in These Factors

| | |
|---|---|
| Evaluation of teaching conducted with video lessons subscale | |
| Factor 1: Positive attitude | |
| videolesson9 | Even if I mark the wrong answer in interactive questions, I can easily learn the correct solution thanks to the feedback. |
| videolesson10 | I like the feedback in the interactive questions to find out where I went wrong. |
| videolesson11 | There must be an interactive question in the video lesson. |
| videolesson12 | I wonder where I went wrong in interactive questions. |
| videolesson13 | Seeing where I went wrong in interactive questions makes it easier for me to understand the topic. |
| videolesson14 | I like to see an encouraging statement that my answer to the interactive questions is correct. |
| videolesson16 | Feedback is necessary in interactive questions. |
| videolesson22 | When I do the activities at the end of the video lesson, I also realize which subjects I lack knowledge about. |
| videolesson23 | I would like to see my name at the top of the points ranking of the activities at the end of the video lesson. |
| videolesson25 | The activities at the end of the video lesson help me to reinforce the topic. |
| videolesson26 | The activities at the end of the video lesson increase the retention of my learning. |
| Factor 2: Negative attitude | |
| videolesson18 | When I cannot solve a question in interactive questions, I stop watching the video lesson. |
| videolesson19 | Learning the subject from video lessons is boring. |
| videolesson30 | I do not continue the activity at the end of the video lesson when I cannot meet the time limit. |
| videolesson35 | I only watch video lessons because my teacher gives me homework. |
| Factor 3: Self-efficacy | |
| videolesson33 | I always search for the correct solutions to the questions I cannot solve in the activities at the end of the video lesson. |
| videolesson34 | I watched the video lesson several times before coming to the face-to-face lesson. |
| videolesson37 | I can watch the video lessons over and over again without getting bored. |
| Factor 4: Advantages | |
| videolesson38 | It allows me to determine the most appropriate time for my learning environment. |
| videolesson39 | It increases my level of active participation in the lesson. |
| videolesson40 | It allows me to manage my own learning process. |
| Factor 5: Accessibility | |
| videolesson2 | I can watch video lessons wherever I want (at home, outside, on the go, etc.). |
| videolesson3 | I can watch video lessons whenever I want. |
| Factor 6: Safety | |
| videolesson45 | I do not find it safe to use unfamiliar learning platforms. |
| videolesson47 | I do not want to provide my personal information to learning platforms, even for educational purposes. |

**Table 6** The Factors and Items in the ETCFFL Subscale and the Items in These Factors

| | |
|---|---|
| Evaluation of teaching conducted with face-to-face lessons subscale | |
| Factor 1: Positive Attitude | |
| facetofacelesson1 | At the beginning of a face-to-face lesson, my teacher summarizes the topic, which helps me to remember what I have learned. |
| facetofacelesson5 | I learn better when I attend a face-to-face lesson by watching a video lesson. |
| facetofacelesson6 | I feel more confident when I attend a face-to-face lesson by watching a video lesson. |
| facetofacelesson7 | I would like to participate more in face-to-face lessons as I have knowledge about the subject by watching videos. |
| facetofacelesson10 | I would like to be the first person to answer questions when I attend a face-to-face class by watching video lessons. |
| Factor 2: Negative Attitude | |
| facetofacelesson2 | It is a waste of time for my teacher to summarize what I have learned in video lessons in face-to-face lessons. |
| facetofacelesson3 | I get bored when my teacher summarizes the topic in video lessons at the beginning of a face-to-face lesson. |

When the items in the final version of the FCME scale that emerged as a result of the explanatory and confirmatory factor analyses were examined, it was found that 24 items were positive and 8 items were negative. Items 10, 11, 16, 19, 24 and 25 in the ETCVL subscale and items 2 and 3 in the ETCFFL subscale are negative and the remaining items are positive (Appendix 1).

***Reliability Analyses***

Item analysis, the differences between the item mean scores of the lower 27% and upper 27% groups formed according to the total scores of the scale were examined using an unrelated t-test. If the differences observed between the groups are significant, it can be interpreted that the internal consistency of the scale is ensured. According to the results of item analysis, it shows to what extent the items distinguish individuals in terms of the measured behavior (Büyüköztürk, 2019).

The FCME scale consists of two different scales, namely the evaluation of teaching conducted with video lessons (ETCVL) subscale and the evaluation of teaching conducted with face-to-face lessons (ETCFFL) subscale. Reliability analyses were therefore conducted separately for both subscales.

With the analyses performed here, a) t-test for the significance of the differences between the item mean scores of the lower 27% and upper 27% groups to be formed according to the scale total scores, b) the reliability of the scale items using item-total correlations, and c) the reliability of the scale using Cronbach's Alpha.

**Table 7** Item Analysis Results of the Items in the Subscales of the ETCVL and ETCFFL

| Item no | Item-total correlation[1] | t (lower %27 – upper %27)[2] |
|---|---|---|
| videolesson2 | .246 | -8.010 |
| videolesson3 | .235 | -7.290 |
| videolesson9 | .497 | -.511 |
| videolesson10 | .519 | -15.227 |
| videolesson11 | .535 | -15.633 |
| videolesson12 | .572 | -16.301 |
| videolesson13 | .537 | -15.158 |
| videolesson14 | .546 | -16.678 |
| videolesson16 | .488 | -13.481 |
| videolesson18 | .323 | -13.725 |
| videolesson19 | .368 | -15.832 |
| videolesson22 | .580 | -16.418 |
| videolesson23 | .538 | -16.647 |
| videolesson25 | .646 | -19.164 |
| videolesson26 | .638 | -19.356 |
| videolesson30 | .312 | -12.564 |
| videolesson35 | .344 | -14.850 |
| videolesson37 | .409 | -13.868 |
| videolesson34 | .458 | -14.692 |
| videolesson40 | .325 | -9.304 |
| videolesson38 | .391 | -11.386 |
| videolesson39 | .352 | -10.168 |
| videolesson45 | .064 | -4.322 |
| videolesson47 | .021 | -3.515 |
| videolesson33 | .516 | -16.777 |
| facetofacelesson1 | .518 | -15.376 |
| facetofacelesson2 | .322 | -14.064 |
| facetofacelesson3 | .391 | -15.492 |
| facetofacelesson5 | .558 | -17.650 |
| facetofacelesson6 | .596 | -18.707 |
| facetofacelesson7 | .624 | -19.656 |
| facetofacelesson10 | .496 | -15.630 |

[1] $n = 871$,
[2] $n_1 = n_2 = 236$,
***$p < .001$

Table 7 shows that the item-total correlations for all items in the ETCVL subscale ranged between .021 and .646 and the t-values were significant (p<.001), while the item-total correlations for all items in the ETCFFL subscale ranged between .322 and .624 and the t-values were significant (p<.001). These results suggest that the reliability of the items in the

ETCVL and ETCFFL subscales is high and that they are discriminative and intended to measure the same behavior.

Table 8 shows the Cronbach’s alpha and McDonald's Omega reliability coefficients of the subscales of the ETCVL and ETCFFL.

**Table 8** Cronbach’s Alpha and McDonald's Omega Values of the FCME Scale and Subscales

| FCME scale | Cronbach’s Alpha reliability coefficient | McDonald's Omega reliability coefficient |
|---|---|---|
| ETCVL subscale | .846 | .840 |
| ETCFFL subscale | .775 | .735 |
| Total | .886 | .879 |

Regarding the reliability of the FCME scale, Cronbach’s alpha coefficient was found to be .846 for the ETCVL subscale, .775 for the ETCFFL subscale, and .886 for the total scale. Since these values are greater than .70 (Büyüköztürk, 2019), it was found that the FCME scale is reliable (Table 8). In addition, since McDonald's Omega (ω) values are above .70, it is understood that the reliability of the scale is sufficient and structural reliability is also provided (McDonald, 1985).

Table 9 shows Cronbach’s alpha and McDonald's Omega reliability coefficients of the factors belonging to the ETCVL subscale and the ETCFFL subscale of the FCME scale.

**Table 9** Cronbach’s Alpha and McDonald's Omega Values of the Subscales and Factors of the Subscales

| Subscales | Factors | Cronbach’s Alpha reliability coefficient | McDonald's Omega reliability coefficient |
|---|---|---|---|
| Evaluation of teaching conducted with video lessons (ETCVL) | Positive attitude | .851 | .888 |
| | Negative attitude | .751 | .752 |
| | Self-efficacy | .740 | .752 |
| | Advantage | .630 | .630 |
| | Accessibility* | .548 | - |
| | Safety* | .555 | - |
| Evaluation of teaching conducted with face-to-face lessons (ETCFFL) | Positive attitude | .813 | .816 |
| | Negative attitude* | .804 | - |

* Omega cannot be estimated because the number of items is less than 3.

The Cronbach’s alpha reliability coefficients of the positive attitude, negative attitude, and self-efficacy factors of the ETCVL subscale were found to be .851, .751, and .740, respectively. In addition, McDonald's Omega (ω) values for the factors of the subscales ranged between .630 and .888.

Since these values were between .70 and .90, these three factors were accepted as reliable. The Cronbach's alpha reliability coefficient of the advantage factor was found to be .630, which is acceptable as it is between .60 and .70. The Cronbach's alpha reliability coefficients of the accessibility and safety factors were calculated as .548 and .555, respectively. Since the reliability coefficients of the reliability factor are between .50 and .60, it is known that the reliability of this factor is weak. However, since the items in this factor are important and necessary for the evaluation of the teaching according to the FC model, it was deemed appropriate to keep it as it is.

The Cronbach's alpha reliability coefficients of the positive attitude and negative attitude factors of the ETCFFL subscale were calculated as .813 and .804, respectively. Since these values were between .70 and .90, these two factors were accepted as reliable.

## Conclusion, Discussion, and Suggestions

This study aims to develop a scale for middle school students' evaluation of mathematics teaching according to the FC model. According to the FC model, students are taught the subject with a video lesson before coming to class, and during class, the subject is summarized and reinforced with various activities. Therefore, the FCME scale developed within the scope of the research consists of two subscales, namely the "Evaluation of Teaching Conducted with Video Lessons" (ETCVL) subscale and the "Evaluation of Teaching Conducted with Face-to-Face Lessons" (ETCFFL) subscale

In the first stage of the scale development process, a literature review was conducted. Within the scope of the information obtained, mathematics teaching was designed according to the FC model, and a sample application was made to middle school students. The opinions of the students about the model were obtained by interviewing the volunteer students who took part in the implementation. An item pool consisting of 90 items was obtained in line with the opinions of the students. This item pool was sent to 11 experts and expert opinions were received. Edits were made by taking expert opinions into consideration and a draft scale consisting of 64 items was obtained. A pilot study was conducted with this draft scale. Validity and reliability analyses were conducted according to the data obtained after the application and a final 5-point Likert-type scale consisting of 32 items in total was obtained.

Şahin and Boztunç Öztürk (2018) examined the scale development studies according to the "group size" variable and found that only 12 articles out of 69 articles had group sizes in the range of 500-999. Based on this result, it can be said that this study, in which 871 students participated, is one of the few scale development studies.

Şahin and Boztunç Öztürk (2018) examined scale development studies according to the "ratio of participants per item". The number of articles with a ratio of 20 and above is only 4 out of 69 articles. In this study, the ratio of participants per item was 27.22 (871:32), making it one of the few scale development studies in the literature. Similarly, the number of articles using "exploratory factor analysis, confirmatory factor analysis" among the methods of determining construct validity is 32 out of 69 articles. 19 articles use sub-upper group analysis and item-test correlation in item analysis and 45 articles examine internal consistency in determining reliability.

In this scale development study, internal consistency was examined to determine reliability, and Cronbach alpha values were calculated. According to the results of EFA and CFA, the ETCVL subscale consists of 25 items and a 6-factor structure. The factors in the ETCVL subscale are named positive attitude, negative attitude, self-efficacy, advantage, accessibility, and safety when analyzed in terms of the items they contain. The reliability coefficient of the ETCVL subscale was calculated as .846 and it was concluded that it was reliable. When the reliability coefficients of the factors were calculated,

- The positive attitude factor consists of 11 items and its reliability coefficient was found to be .851,
- The negative attitude factor consists of 4 items and its reliability coefficient was found to be .751,
- The self-efficacy factor consists of 3 items and its reliability coefficient was found to be .740,
- The advantage factor consists of 3 items and its reliability coefficient was found to be .630,
- The accessibility factor consists of 2 items and its reliability coefficient was found to be .548 and
- The safety factor consists of 2 items and the reliability coefficient was found to be .555.

The reliability coefficients of the advantage, accessibility, and safety factors were lower than .70. However, it was decided to keep these items and factors in the scale because the items in these factors were considered to be necessary for the scale, they could not be gathered under a single factor in terms of the characteristics they measured, and the final version of the ETCVL subscale was calculated to be reliable.

According to the results of EFA and CFA, the ETCFFL subscale consists of 7 items and a 2-factor structure. When the common characteristics of the items that make up the factors in the ETCFFL subscale were examined, these factors were named positive attitude and negative attitude.

The reliability coefficient of the ETCFFL subscale was calculated as .775 and it was concluded that it was reliable. When the reliability coefficients of the factors are calculated;

- The positive attitude factor consists of 5 items and its reliability coefficient was found to be .813, and
- The negative attitude factor consists of 2 items and the reliability coefficient was found to be .804.

After calculating the reliability coefficients of the two subscales of the FCME scale, the reliability coefficient of the scale as a whole was found to be .886 and it was concluded that the FCME scale was reliable.

There are 24 positive and 8 negative items in the FCME scale. The scores that can be obtained from the scale vary between 32 and 160.

In the literature review, four scales related to the FC model were found. Of these scales, two are scale development studies (Akgün, 2015; Erensayın, 2019), one is the adaptation of an existing scale in the literature into Turkish (Durak, 2017), and the other is the adaptation of a scale for students in the literature for teachers (Kurtoğlu, 2019). When these four studies were evaluated in terms of disciplines, it was determined that they were not in the field of mathematics. Within this scope, no scale was found in the literature for the evaluation of the use of the FC model in mathematics teaching. In conclusion, a valid and reliable scale for the evaluation of the use of the FC model in mathematics teaching was developed in this study.

The validity and reliability analyses of the scales developed by Erensayın (2019), Kurtoğlu (2019), and Durak (2017) are available. Nevertheless, it was found that the scale developed by Akgün (2015) did not provide information on validity and reliability analyses. Detailed information on the reliability and validity analyses of the FCME scale was given and the scale was found to be reliable and valid.

Akgün (2015) developed a scale titled “Opinion Questionnaire on flipped classrooms”. However, it is important that the study did not include validity and reliability studies. Furthermore, the scale items evaluated the “inverted-flat classroom practice” in general, which negatively affects the content validity of the scale. In this study, unlike Akgün (2015), there are more detailed items for the evaluation of teaching and items that allow the evaluation of every dimension of teaching. Another difference is that in this study, the two phases of the teaching of the FC model (teaching with video lessons and teaching with face-to-face lessons) were handled separately. Studies also differ in terms of the online platforms used in the teaching phase with video lessons.

The scale items developed in Erensayın’s (2019) study and the scale items developed in this study have similar items. The difference between this study and Erensayın’s (2019) study is that there is a negative attitude sub-dimension belonging to both stages of the instruction.

It was determined that the study conducted by Durak (2017) and this study were similar in terms of items. The difference between this study and Durak’s (2017) study is that there are negative items in the scale developed and it is a scale for the evaluation of two important stages of teaching.

When the scale development studies were evaluated in terms of the study group, they were conducted with secondary school teachers from different branches (Erensayın, 2019), middle school teachers from different branches, and middle school (5th and 8th grade) students (Kurtoğlu, 2019), middle school 5th and 6th-grade students (Durak, 2017), and middle school 5th-grade students (Akgün, 2015). This study was conducted with all grade levels (5th, 6th, 7th, and 8th grade students). In this way, a scale was created for all grade levels at the secondary school level.

Erensayın (2019) used exploratory factor analysis in the scale development study and Durak (2017) used confirmatory and exploratory factor analysis in the study. In the study conducted by Durak (2017), although it is not specified which program was used, it is seen that the visuals and fit criteria of the Structural Equation Model were used for the confirmatory factor. In this direction, it can be said that the scale of LCME, whose validity was increased with the AMOS statistical program compatible with the Structural Equation Model, was introduced to the field.

In conclusion of this study, suggestions for researchers are given below.

- This scale can be used to evaluate the teaching of mathematics according to the FC model.
- The scale can also be used for primary school, high school, and university students by conducting a reliability study.
- Since the items in the scale are general statements, they can be used for the evaluation of instruction for each learning area or sub-learning area of mathematics.
- Although the FCME scale was prepared for the mathematics course, it does not contain mathematical terms, so it can be used in the evaluation of teaching based on the FC model for different branches.
- This scale can be used in the studies to be conducted by teachers, lecturers, and prospective teachers on education.

**Compliance with Ethical Standards**

*Disclosure of potential conflicts of interest*

There was no conflict of interest in this study.

*Funding*

No financial support was received.

*CRediT author statement*

This study is a part of the master's thesis of the second author under the supervision of the first author.

*Research involving Human Participants and/or Animals*

This research was conducted with the permission obtained in accordance with the decision taken at the meeting of the Science and Engineering Sciences Ethics Committee dated 20.04.2021 and numbered 2021/2

## Matematik Eğitiminde Ters Yüz Edilmiş Sınıf Modelinin Kullanımına Yönelik Bir Ölçek Geliştirme Çalışması

**Özet:**

Bu araştırmada ortaokul matematik öğretiminde Ters Yüz Edilmiş Sınıf Modelinin kullanımının değerlendirilmesini sağlayacak geçerli ve güvenilir bir ölçek geliştirmek amaçlanmıştır. Araştırmanın deseni tarama deseni olarak belirlenmiştir. Ölçeğin geliştirilme sürecinde ilk olarak madde havuzu oluşturulmuş ve uzman görüşü alınmıştır. Uzmanların görüşlerine göre düzenlemeler yapılmıştır. Çalışma grubu basit seçkisiz örneklemeyle belirlenmiş olup 5,6,7 ve 8.sınıflardan olmak üzere toplam 871 ortaokul öğrenciyle uygulama gerçekleştirilmiştir. Elde edilen verilere göre ölçeğin geçerlik ve güvenirlik analizleri SPSS ve AMOS programlarıyla yapılmıştır. Ters Yüz Edilmiş Sınıf Modeli Değerlendirme(TYESMD) ölçeği, "Video Ders İle Yapılan Öğretimin Değerlendirilmesi(VDYÖD) Alt Ölçeği" ve "Yüz Yüze Ders İle Yapılan Öğretimin Değerlendirilmesi(YDYÖD) Alt Ölçeği" olmak üzere iki alt ölçekten oluşmaktadır. Yapılan analizler sonucunda toplamda 32 maddeden oluşan TYESMD ölçeği geliştirilmiştir. VDYÖD alt ölçeğinin güvenirlik katsayısı .846, YDYÖD alt ölçeğinin .775 ve bir bütün olarak ölçeğin güvenirlik katsayısı .886 olarak bulunmuştur. Yapılan analizler sonucunda TEYSMD Ölçeğinin geçerli ve güvenilir bir ölçek olduğu tespit edilmiştir.

*Anahtar kelimeler*: Ters yüz edilmiş sınıf modeli, matematik eğitimi, ortaokul öğrencileri, ölçek geliştirme.

## References


Akgün, M. (2015). *The effect of flipped classroom on students' academic achievement and views* (Publication No. 423423) [Master's thesis, Fırat University]. Council of Higher Education National Thesis Center.

Altunışık, R., Coşkun, R., Bayraktaroğlu, S., & Yıldırım, E. (2005). *Sosyal bilimlerde araştırma yöntemleri*: Spss *uygulamalı* (4th ed.). Sakarya.

Arslan, U. (2021). *Investigation of the effect the flipped classroom model on middle school students' academic achievement and self-regulation skills in mathematics course* (Publication No. 695917) [Master's thesis, Çukurova University]. Council of Higher Education National Thesis Center.

Atal, D., & Koçak Usluel, Y. (2011). Elementary school students' use of technology in and out of school. *Hacettepe University Journal of Education*, 41, 24-35. https://dergipark.org.tr/en/download/article-file/87383

Baker, J. W. (2000). The "classroom flip": Using web course management tools to become the guide by the side. In: Chambers JA (ed.), *Selected Papers from the 11th International Conference on College Teaching and Learning (*pp. 9–17*)*. Florida Community College at Jacksonville. https://upcea.edu/wp-content/uploads/2020/09/The-Classroom-Flip-Baker.pdf

Balkaya, M. (2022). *Developing and implementing a measurement tool related to Youtube use with learning purposes for middle school student* (Publication No. 731719) [Master's thesis, Çanakkale Onsekiz Mart University]. Council of Higher Education National Thesis Center.

Bender, W.N. (2018). Ters yüz edilmiş stem sınıfı. In S. Durmuş, A. S. İpek, & B. Yıldız (Trans. ed.)*, STEM öğretimi için 20 strateji* (1st ed., pp. 73-82). Nobel.

Bergmann, J. (2011, May 6). The history of the flipped class. *Flippedclass.com* https://flippedclass.com/the-history-of-the-flipped-class/

Bolatlı, Z. (2018). *Determining the academic achievement of students who use flipped classroom method supported by mobile application and their views on cooperative learning* (Publication No. 534559) [Master's thesis, Selçuk University]. Council of Higher Education National Thesis Center.

Bulut, R. (2019). *Investigation of the use of the flipped clasroom model in teaching the ratio-proportion* (Publication No. 563939) [Master's thesis, Erzincan Binali Yıldırım University]. Council of Higher Education National Thesis Center.

Büyüköztürk, Ş. (2019). *Sosyal bilimler için veri analizi el kitabı istatistik, araştırma deseni spss uygulamaları ve yorum* (25th ed.). Pegem.

Büyüköztürk, Ş., Kılıç Çakmak, E., Akgün, Ö.E., Karadeniz Ş., & Demirel, F. (2018). *Bilimsel araştırma yöntemleri* (24th ed.). Pegem.

Creswell, J. W. (2017). *Araştırma deseni nitel, nicel ve karma yöntem yaklaşımları* (3rd ed.). (S.B., Demir, Trans.). Eğiten Kitap.

Demir, B. (2016). *A scale development for anxiety of student for finding work with structural equation model: An application faculty of economics and administrative sciences in Cumhuriyet University* (Publication No. 459470) [Master's thesis, Cumhuriyet University]. Council of Higher Education National Thesis Center.

Durak, H. Y. (2017). Turkish adaptation of the flipped learning readiness scale for middle school students. *Bartın University Journal of Faculty of Education, (6)*3, 1056-1068. https://doi.org/10.14686/buefad.328826

Erensayın, E. (2019). *Examination of the applicability of the flipped classroom model in secondary education according to teacher perceptions* (Publication No. 580481) [Master's thesis, Van Yüzüncü Yıl University]. Council of Higher Education National Thesis Center.

Güç, F. (2017). *The effect of the flipped classroom practice on the rational numbers and operations with rational numbers* (Publication No. 478696) [Master's thesis, Amasya University]. Council of Higher Education National Thesis Center.

Hao, Y. (2016). Middle school students' flipped learning readiness in foreign language classrooms: Exploring its relationship with personal characteristics and individual circumstances. *Computers in Human Behavior*, *59,* 295-303. https://doi.org/10.1016/j.chb.2016.01.031

Houston, K. (2020). *The flipped classroom's impact on math anxiety, self-efficacy, and motivation at the middle school level* (Publication No. 28260171) [Doctoral dissertation, Cabrini University]. ProQuest Dissertations & Theses Global.

Karagöz, Y. (2021). *SPSS AMOS META uygulamalı biyoistatistik* (3rd ed.). Nobel.

Koehler, M., & Mishra, P. (2009). What is technological pedagogical content knowledge (Tpack)?. *Contemporary Issues in Technology and Teacher Education, 9*(1), 60-70. https://www.learntechlib.org/primary/p/29544/

Kocaman Karoğlu, A. (2015). The changing nature of storytelling by means of technology in the instructional process: digital storytelling. *Educational Technology Theory and Practice*, *5*(2), 89-106. https://doi.org/10.17943/etku.29277

Kurtoğlu, C. (2019). *Investigation of the readiness of middle school teachers and students regarding the flipped classroom model* (Publication No. 605297) [Master's thesis, Fırat University]. Council of Higher Education National Thesis Center.

McDonald, R. P. (1985). *Factor analysis and related methods* (1st ed.). Erlbaum. https://doi.org/10.4324/9781315802510

MoNE (2018). *Matematik dersi öğretim programı (ilkokul ve ortaokul 1, 2, 3, 4, 5, 6, 7 ve 8. sınıflar)*. http://mufredat.meb.gov.tr/ProgramDetay.aspx?PID=329

Mutluoğlu, A., & Erdoğan, A. (2016). Examining primary mathematics teachers' technological pedagogical content knowledge (TPACK) levels according to their preferred teaching styles. *OPUS International Journal of Society Researches*, *6*(10), 102-126. https://dergipark.org.tr/en/pub/opus/issue/22701/242324#article_cite

Nanto, Z. (2021). *The relationship between error management culture, work engagement and organizational creativity* (Publication No. 659550) [Doctoral dissertation, Fırat University]. Council of Higher Education National Thesis Center.

Özdemir, A. (2016). *Flipped classroom model practice focused on blended learning in secondary mathematics education* (Publication No. 429471) [Doctoral dissertation, Fırat University]. Council of Higher Education National Thesis Center.

Şahin, M. G., & Boztunç Öztürk, N. (2018). Scale Development Process in Educational Field: A Content Analysis Research. *Kastamonu Education Journal*, 26(1), 191-199. https://doi.org/10.24106/kefdergi.375863

Wei, X., Cheng, I., Chen, N. S., Yang, X., Liu, Y., Dong, Y., & Zhai, X. (2020). Effect of the flipped classroom on the mathematics performance of middle school students. *Educational Technology Research and Development*, *68*, 1461-1484. https://doi.org/10.1007/s11423-020-09752-x

Yurdugül, H. (2005, September 28-30). *Ölçek geliştirme çalışmalarında kapsam geçerliği için kapsam geçerlik indekslerinin kullanılması*. XIV. Ulusal Eğitim Bilimleri Kongresi, Pamukkale Üniversitesi Eğitim Fakültesi, Denizli. http://yunus.hacettepe.edu.tr/~yurdugul/3/indir/PamukkaleBildiri.pdf

Yurdugül, H., & Bayrak, F. (2012). Content validity measures in scale development studies: Comparison of content validity index and kappa statics. *Hacettepe University Journal of Education*, Special Issue 2, 264-271. http://efdergi.hacettepe.edu.tr/yonetim/icerik/makaleler/655-published.pdf

## Appendix 1

*The Flipped Classroom Model Evaluation (FCME) Scale*

| | **Evaluation of Teaching Conducted with Video Lessons (ETCVL) Subscale** | **Strongly Disagree** | **Disagree** | **Neither agree nor disagree** | **Agree** | **Strongly Agree** |
|---|---|---|---|---|---|---|
| **1** | I can watch video lessons wherever I want (at home, outside, on the go, etc.). | | | | | |
| **2** | I can watch video lessons whenever I want. | | | | | |
| **3** | Even if I mark the wrong answer in interactive questions, I can easily learn the correct solution thanks to the feedback. | | | | | |
| **4** | I like the feedback in the interactive questions to find out where I went wrong. | | | | | |
| **5** | There must be an interactive question in the video lesson. | | | | | |
| **6** | I wonder where I went wrong in interactive questions. | | | | | |
| **7** | Seeing where I went wrong in interactive questions makes it easier for me to understand the topic. | | | | | |
| **8** | I like to see an encouraging statement that my answer to the interactive questions is correct. | | | | | |
| **9** | Feedback is necessary in interactive questions. | | | | | |
| **10** | When I cannot solve a question in interactive questions, I stop watching the video lesson. | | | | | |
| **11** | Learning the subject from video lessons is boring. | | | | | |
| **12** | When I do the activities at the end of the video lesson, I also realize which subjects I lack knowledge about. | | | | | |
| **13** | I would like to see my name at the top of the points ranking of the activities at the end of the video lesson. | | | | | |
| **14** | The activities at the end of the video lesson help me to reinforce the topic. | | | | | |
| **15** | The activities at the end of the video lesson increase the retention of my learning. | | | | | |
| **16** | I do not continue the activity at the end of the video lesson when I cannot meet the time limit. | | | | | |
| **17** | I always search for the correct solutions to the questions I cannot solve in the activities at the end of the video lesson. | | | | | |
| **18** | I watched the video lesson several times before coming to the face-to-face lesson. | | | | | |
| **19** | I only watch video lessons because my teacher gives me homework. | | | | | |

(continued)

| | **Evaluation of Teaching Conducted with Video Lessons (ETCVL) subscale** | **Strongly Disagree** | **Disagree** | **Neither agree nor disagree** | **Agree** | **Strongly Agree** |
|---|---|---|---|---|---|---|
| **20** | I can watch the video lessons over and over again without getting bored. | | | | | |
| **21** | It allows me to determine the most appropriate time for my learning environment. | | | | | |
| **22** | It increases my level of active participation in the lesson. | | | | | |
| **23** | It allows me to manage my own learning process. | | | | | |
| **24** | I do not find it safe to use unfamiliar learning platforms. | | | | | |
| **25** | I do not want to provide my personal information to learning platforms, even for educational purposes. | | | | | |

| | **Evaluation of Teaching Conducted with Face-to-Face Lessons (ETCFFL) Subscale** | **Strongly Disagree** | **Disagree** | **Neither agree nor disagree** | **Agree** | **Strongly Agree** |
|---|---|---|---|---|---|---|
| **1** | At the beginning of a face-to-face lesson, my teacher summarizes the topic, which helps me to remember what I have learned. | | | | | |
| **2** | It is a waste of time for my teacher to summarize what I have learned in video lessons in face-to-face lessons. | | | | | |
| **3** | I get bored when my teacher summarizes the topic in video lessons at the beginning of a face-to-face lesson. | | | | | |
| **4** | I learn better when I attend a face-to-face lesson by watching a video lesson. | | | | | |
| **5** | I feel more confident when I attend a face-to-face lesson by watching a video lesson. | | | | | |
| **6** | I would like to participate more in face-to-face lessons as I have knowledge about the subject by watching videos. | | | | | |
| **7** | I would like to be the first person to answer questions when I attend a face-to-face class by watching video lessons. | | | | | |