Education and Science

Eğitim ve Bilim

April: 2026 Vol: 51 Issue: 226 / Nisan: 2026 Cilt: 51 Sayı: 226

ISSN 1300-1337

Education and Science
Eğitim ve Bilim

1300-1337

Vol 51, Issue 226, 2026

Eğitim ve Bilim
Education and Science

Owner on behalf of the Turkish Education Association (TEA) Economic Enterprises

Sinan Selçuk PEHLİVANOĞLU

Education and Science is official journal of Turkish Education Association. Since Education and Science is a peer-reviewed journal, only research articles are given place. Manuscripts submitted to the journal, primarily reviewed by the editors and associate editors, are forwarded to 2 reviewers specialized in related content areas for more detailed evaluation. Education and Science Journal is published quarterly. Education and Science Journal strives for increasing number of academic studies that offer concrete solutions for various areas of education that requires improvements, for all levels of education starting from pre-school up to university level, and that provide support for improvements in the field of education, and that could make contribution to professional improvement of teachers and academicians.

Education and Science Online Access Web:
http://egitimvebilim.ted.org.tr
e-mail: egitimvebilim@ted.org.tr

Publish Date: April 2026

Education and Science is indexed in Social Sciences Citation Index (SSCI), Scopus, ProQuest, EBSCO Host, and TR Dizin.

Publication Type: Quarterly Published Academic Journal

Adress: Turkish Education Association (TEA) Economic Enterprises **(**Türk Eğitim Derneği İktisadi İşletmesi), TED Ankara Koleji Vakfı Okulları İncek Kampüsü Taşpınar Mah. 2800. Cadde No: 5 F Blok İncek - Gölbaşı / Ankara, Türkiye

Phone: +90 (312) 939 50 82 **Fax:** +90 (312) 417 23 74

Print: Teknoart Digital Ofset Reklamcılık Matbaacılık İthalat İhracat Sanayi ve Ticaret Ltd. Şti.

Adress: Cevizlidere Mah. 1288. Sokak No:1/1 Balgat, Çankaya/Ankara, Türkiye

Certificate No: 47644 | **Phone:** +90 (312) 473 92 97 | **e-mail:** info@printandsmile.com.tr

Education and Science

Vol 51 (2026) No 226

Contents

Education and Science

Original Article

Vol 51 (2026) No 226 265-300

"I think, therefore i draw": an alternative action research study on P4C practices

Seval Çiğdemir [1]

Abstract

This study is structured around the use of drawings by first-grade primary school students to express their thoughts during Philosophy for Children (P4C) practices. For this purpose, preparatory activities were conducted through drama; selected picture books were used as stimuli; themes were identified for each book through community of inquiry sessions; and, finally, all students were enabled to express their ideas through drawings. Designed as an action research study within the framework of qualitative research methods, the research group consisted of 35 first-grade students. Data were collected using a checklist, interview forms, a teacher journal, audio recordings, and field notes. As a result of the practices conducted over two months, it was observed that philosophical inquiry through students' own drawings was feasible for younger learners who experience difficulty in verbal expression. The findings revealed that students who were supported in expressing their thoughts demonstrated not only improved thinking skills but also strengthened social skills and reduced communication difficulties. Moreover, an overall increase in students' capacity to respond and engage in classroom dialogue was observed. Other areas of development observed included generating opposing viewpoints, producing new ideas, and gaining flexibility by changing opinions when necessary. Future studies are recommended to explore alternative ways of including students with limited expressive language skills in P4C activities.

Keywords

P4C

Picture books

Drawing

Action research

Expression of thought

Article Info

Received: 07.18.2024

Accepted: 02.02.2026

Published Online: 04.30.2026

DOI: 10.15390/ES.2026.2510

Introduction

Primary education curricula aim to equip individuals with the capacity for problem-solving, entrepreneurship, and responsibility, and to meet the demands of the contemporary world (Ministry of National Education [MoNE], 2018). The concept of "teaching," which dominated the previous century, has largely been replaced by "learning" (Aydın, 1998). As learners take an active role in knowledge acquisition, they are expected to develop skills such as innovative, critical, and creative thinking (Haywood, 2020; Özlem, 1997). In response to this need, the Turkish Ministry of National Education established goals within the framework of the 2023 Education Vision to support the development of individuals who can solve problems and generate new ideas from an early age (MoNE, 2020). In this

[1] Ministry of National Education, Ankara, Türkiye, sevalcgdmr@gmail.com

context, the in-service "Philosophy for Children Basic Training" provided to teachers can be considered a significant initiative.

Philosophy is defined as the investigation of existence and knowledge, doctrine, worldview, and abstract thinking on a particular subject (Turkish Language Association [TDK], 2021). Etymologically, the word philosophy derives from the Greek words *philia* (love) and *sophia* (wisdom), and can be described as the love of knowledge (Eren et al., 1998). In this context, knowledge refers not to the "positive sciences," which aim to explain the universe through distinct disciplines, but rather to forms of understanding that cannot be fully explained or definitively answered (Dirican & Deniz, 2020). Philosophers and children share several common characteristics, such as a sense of wonder and curiosity, as well as a persistent tendency to ask questions. These shared traits naturally give rise to philosophy, which requires thinking, questioning, and discussion (Çiçek, 2017; Droit, 2017; Kantarcı, 2013; Wartenberg, 2018). Since children can think about concepts such as good and evil independently of social pressures and tend to be more open-minded in critical thinking, critical thinking education should begin at an early age (Droit, 2017).

The fundamental questions posed by early philosophers, particularly those concerning the universe and morality, closely resemble the questions children ask as they try to understand the world. For this reason, the ancient period has often been described metaphorically as the "childhood of philosophy" (Önal, 2017). One of the approaches proposed to prevent the loss of children's innate questioning ability—believed to diminish over time—and to support its development is Philosophy for Children (P4C) (Boyacı et al., 2018; Çiçek, 2017; McCall, 2017; Taşdelen, 2013; Wartenberg, 2018). UNESCO (2007) outlines the objectives of philosophy education for preschool and primary school children in six stages:

- Encouraging independent thinking
- Educating reflective individuals
- Supporting children's personal development
- Developing language and communication skills
- Enabling conceptual thinking
- Designing instructional approaches appropriate to children's developmental levels

To achieve these goals through philosophy education, three widely accepted approaches are commonly employed. The first, P4C, focuses on engaging children and adolescents in philosophical inquiry. The second, the "community of philosophical inquiry," targets young people and adults, while the "Socratic method" is applied exclusively with adults (Boyacı et al., 2018; McCall, 2017). P4C (Philosophy for Children), one of the approaches aimed at developing thinking skills, was developed by Matthew Lipman in the 1960s (Akkocaoğlu Çayır, 2015). The acronym P4C is derived from the initial letters of *creative thinking*, *caring thinking*, *critical thinking*, and *collaborative thinking* (Erdoğan, 2018). As its name implies, this method aims not to teach children theoretical philosophy or philosophical concepts, but rather to develop their thinking skills and engage them in philosophical practices through workshops (Gür, 2010). In P4C, children discuss, defend, or critique questions they have generated themselves within an appropriate inquiry environment. Although the method has gained increasing popularity in recent years, its theoretical foundations are grounded in the ideas of Vygotsky, Dewey, and Mead (Lipman, 2003). Numerous studies have demonstrated that P4C activities contribute to the development of children's thinking skills, including creativity and critical thinking (Bülbül Hüner, 2021; Cassidy et al., 2018; Demirtaş et al., 2018; Jones-Teuben, 2013; Kefeli & Kara, 2008; Lukey, 2006; Mehdizadeh et al., 2019).

A philosophical session typically begins with preparatory and relaxation activities, followed by the presentation of a stimulus (e.g., a book, film, or poem). Children are then given time to reflect on the stimulus and generate questions. In the subsequent inquiry phase, students are expected to formulate

their own questions, which are then organized, related to one another, and narrowed down through a selection process. At this stage, the facilitator (teacher, P4C trainer, etc.) guides the inquiry by employing strategies that encourage staying within the context, sharing new perspectives, and fostering discussion and collaboration. Finally, a general summary is provided, and the session is concluded (Karadağ et al., 2017). As Lipman (2003) emphasizes, when the process ends, students should have more questions than they did at the beginning.

When Philosophy for Children activities are implemented in the classroom, the teacher's role shifts to that of a "facilitator." Teachers who traditionally dominate classroom talk for 70-75% of lesson time—asking questions, managing the class, and often avoiding cognitively demanding activities (Pihlgren, 2013)—become facilitators who organize the environment, support student participation, ask questions that require justification, and enhance students' philosophical awareness (Lipman et al., 1980). Facilitators must attend to several key considerations during the inquiry process, including adhering to the inquiry structure, respecting diverse viewpoints, building trust with children, avoiding suggestion, selecting appropriate stimuli, overcoming impasses, and refocusing discussions when necessary (Lipman et al., 1980; Murris, 2008). In Greek mythology, Theseus escapes the labyrinth by following the thread left by Ariadne. Similarly, during philosophical thinking, the mind may wander into side paths and become lost in a conceptual labyrinth. The facilitator's role, like Ariadne's, is to help participants remain oriented and progress in the right direction (Worley, 2010).

An essential component of Philosophy for Children education is the selection of stimuli. The materials used in discussions may include abstract forms such as stories, short texts, and poems, as well as engaging visuals, artworks, situations, or events (Oktar, 2019). The chosen stimulus should be both significant and debatable, allowing for multiple interpretations and responses. When working with children, it should also be engaging and enjoyable (Matthews, 2000). For this reason, the founders of Philosophy for Children preferred stories as primary stimuli (Lipman et al., 1980). Philosophical activities based on stories enable children to develop a range of skills, including text analysis, comprehension, interpretation, speaking, writing, listening, and self-expression (Direk, 2019; Taş & Dikici Sığırtmaç, 2018). For younger age groups, picture books are frequently used as stimuli. Picture books were first introduced into philosophy education by Murris (1992). However, rather than embedding philosophy explicitly within the text, the aim should be for philosophical thinking to emerge naturally through the interaction between the child, the book, and the reader (Özdemir, 2021). When selecting stories, texts that children can relate to, that capture their attention, and that are concise and clear should be preferred (Lipman et al., 1980). Studies on the use of children's literature in Philosophy for Children practices further support this approach (Akkocaoğlu Çayır, 2015; Fisher, 2001; Karasu, 2018; Mazı, 2008; Murris & Thompson, 2016).

The process of making sense of and expressing intellectual activities such as philosophy, science, and art is mediated through language, which serves as a tool for projecting the structure of the mind onto the external world (Çotuksöken, 1994). In philosophical inquiry, while logical connections between ideas are established through mental activity, these ideas are also communicated to others, thereby fostering social interaction (Lipman, 1998). Because Philosophy for Children activities involve stages such as listening to texts, reading, and verbal questioning, they are fundamentally grounded in linguistic practices (Fisher, 2001; Haynes, 2008; Lipman, 1998). A substantial body of research highlights the impact of P4C practices on children's language development (Cassidy & Christie, 2013; Ferreira, 2004; Jenkins & Lyle, 2010; Jones-Teuben, 2013; Murris & Thompson, 2016; Topping & Trickey, 2014; Ventista, 2019). However, children may experience difficulties in verbal expression due to language delays, autism, hearing impairment, language barriers, or traumatic experiences. Even in the absence of physical impairments, factors such as shyness, lack of self-confidence, or limited opportunities can hinder children's ability to express themselves (Çakmak Tolan & Genç, 2021). Consequently, supportive activities are required, particularly for children who are reluctant to express their thoughts verbally. In such cases, drawings may serve as an alternative means of expression.

In recent years, the use of drawings as a data collection tool in social science research and studies focusing on meaning-making through visual literacy has gained increasing importance (Hopperstad, 2010; Mayers, 2011; Thomson, 2008). According to dual coding theory, information is stored in the mind in both verbal and visual forms, and the presentation of recorded information involves a complex structure (Paivio, 1986, as cited in Akyol, 2006). Therefore, when verbal and visual representations support one another, children can express their thoughts more holistically. Research indicates that children can generate powerful and creative ideas through drawings and use them to understand the world and articulate complex concepts (Kendrick & McKay, 2009; Whitfield, 2009). In philosophical practices, the significance of drawings lies not in what the image literally represents, but in the ideas children generate while explaining their drawings. Due to the cognitive structures characteristic of preschool and primary school children, the use of visuals increases the likelihood of accessing children's ideas through visual representations (Clark & Moss, 2011; Collado, 1999; Thomson, 2008) and for children, drawing functions not only as a means of expressing internal experiences but also as a way of making sense of and organizing their relationships with the environment and society (Malchiodi, 1998).

Although the impact of P4C activities on essential developmental outcomes such as thinking skills, communication, socialization, and critical thinking is well documented, no study has been identified that specifically examines how these activities influence the thinking of students who experience difficulties with self-expression. Studies that focus solely on numbers and words tend to overlook the role of drawings in revealing raw emotions and unarticulated experiences (Cox et al., 2014). Accordingly, this study investigates both the general effects of P4C practices using picture books and their specific impact on students who struggle with self-expression. The study seeks to answer the following research questions:

1. What effect does the use of drawings in P4C activities with picture books have on the thinking skills of first-grade students?
2. What effect does the use of drawings in P4C activities with picture books have on first-grade students who experience difficulties in self-expression?

Method

In this study, which examines the process of expressing thoughts through drawings in P4C practices based on picture books, the action research method was employed, with the researcher serving as both practitioner and researcher. Action research is a research approach in which professionals who are directly involved in the process advance the research through cyclical stages rather than remaining external observers (Calhoun, 1994; Corey, 1949; Hinchey, 2008). This design is commonly used by practitioners to address their own problems or to improve existing practices, and it appears in the literature under various labels such as collaborative, contextual, emancipatory, and participatory action research (Baumfield et al., 2013; O'Brien, 2001).

Structurally, the present study falls within the category of **technical/scientific/collaborative action research** as classified by McKernan (2008). In this classification, the researcher—who possesses a strong command of the theoretical framework—tests a new approach and carries out ongoing evaluation and modifications throughout the process (Yıldırım & Şimşek, 2008, p. 296). The cyclical process followed in the present study is illustrated in Figure 1.

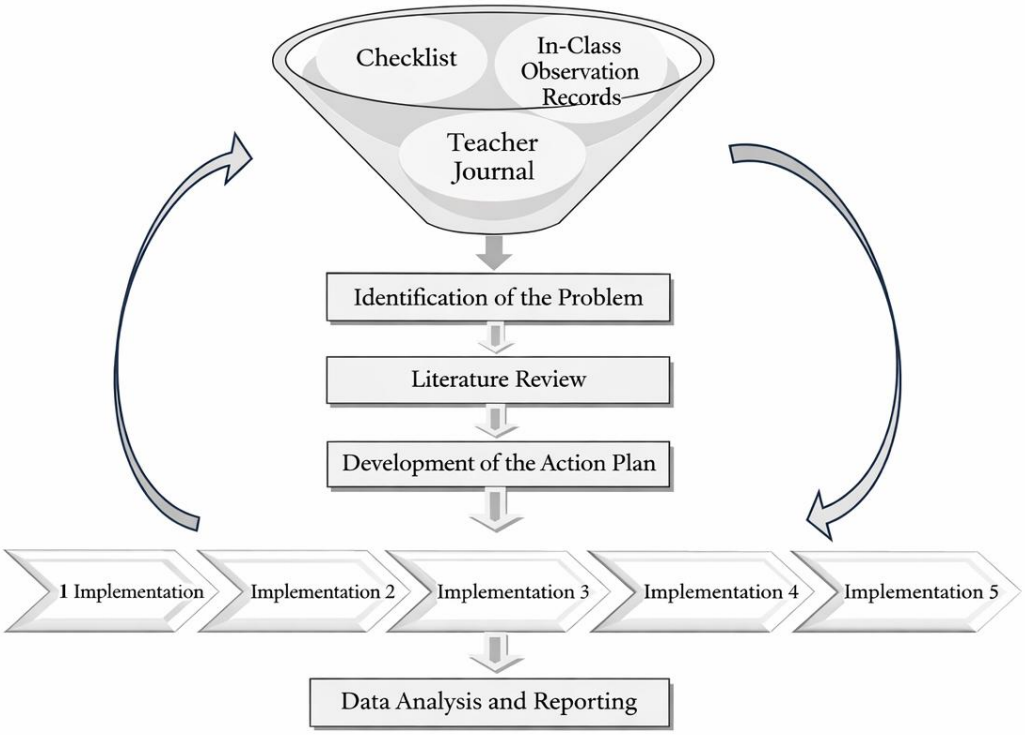


Figure 1. Action research plan

As illustrated in Figure 1, in-class philosophy workshops were designed using P4C practices and picture books. Throughout these workshops, students' participation in the process was examined through teacher journals, audio recordings, and a checklist form. As a result of the analyses, it was identified that some students experienced difficulties in expressing their thoughts. To identify possible solutions to this issue, a literature review was conducted. Subsequently, an action plan was designed to address the identified problem. Observation forms, visual and audio recordings, semi-structured interview forms, and field notes were analyzed progressively throughout the process, enabling shortcomings to be addressed and the implementation to be improved. Finally, all data collected during the research process were organized and reported.

Identification of the Problem

The researcher also served as the classroom teacher in the context where the study was conducted. During implementation, some students were reluctant to participate, and in some cases, only a limited number actively engaged in the inquiry process, while others remained largely passive. To analyze this situation more systematically, the researcher decided to record the final workshop session and to develop a checklist and a teacher journal for data analysis.

The checklist included items related to students' participation, such as taking turns to speak, proposing new ideas, changing opinions during the inquiry process, and developing opposing viewpoints (Appendix 1). Analysis of the checklist revealed that students coded as S4, S8, S18, S21, and S25 demonstrated insufficient participation in the workshops. To both increase participation among the identified students and enhance the effectiveness of P4C workshops for the entire class, a literature review was conducted. Based on the assumption that students might express themselves more comfortably through alternative modes, drawings were adopted as a strategy in subsequent action cycles.

Research Setting and Participants

The participants in the study were 35 first-grade students enrolled in a public primary school. The school reflects the national average of Türkiye in terms of socio-cultural and socio-economic characteristics. According to the Socio-Economic Development Index Report, the province in which the school is located ranks second nationwide, while the district ranks 47th (Republic of Türkiye Ministry of Industry and Technology, 2023). However, due to the neighborhood's heterogeneous structure and exposure to internal migration, the local context shows socio-economic characteristics below the average. Established in 2021, the school currently comprises 32 classrooms and serves 1,033 students. Class sizes range from 30 to 39 students. The school is equipped with a music and art workshop, an indoor sports hall, and a library.

The classroom in which the study was conducted included 35 first-grade students, 18 girls, and 17 boys. Among these students were two mainstreamed students formally diagnosed with *specific learning difficulties* and *mild intellectual disability*, respectively. In addition, two students diagnosed with anxiety disorders and receiving ongoing therapy were present in the classroom. At the time the study commenced, all students except one mainstreamed student had acquired independent reading and writing skills. The classroom was equipped with a smartboard and a classroom library of books that students could access freely at any time.

The study was designed with a dual focus. In addition to monitoring the development of thinking skills across the entire class, the study also examined in depth the progress of five students who experienced difficulties in self-expression. Detailed information regarding these students with self-expression difficulties is presented in the following section.

Table 1. Characteristics of students included in the individual inquiry

Student	Academic Achievement	Social Relationships	Family Background	Participation in Other Lessons
S4	High academic achievement	Socially withdrawn; diagnosed with an anxiety disorder; interacts with a limited number of peers	Supportive family; maintains effective communication with family members	Consistently willing to participate in lessons
S8	Moderate academic achievement	Maintains typical peer relationships	Youngest child in the family; frequently compared with older siblings; experiences familial pressure	Generally willing to participate in lessons
S18	Moderate academic achievement	Limited positive peer interactions; highly active; occasionally displays aggressive behaviors	Family is not open to communication and perceives behaviors as typical	Participates intermittently; becomes silent when explicitly encouraged to speak
S21	Low academic achievement	Maintains positive peer relationships	Family is open to communication	Rarely takes turns to speak; low level of participation
S25	Moderate academic achievement	Harmonious peer relationships; generally quiet	Family is open to communication	Willing to participate when the topic is familiar

According to Table 1, student S4 demonstrates high academic achievement; students S8, S18, and S25 exhibit moderate academic achievement; and student S21 exhibits low academic achievement. Including students from different academic levels broadens the scope of the study. In terms of peer relationships, S8 and S21 display typical social behaviors, whereas S18 occasionally engages in verbal and physical aggression. Students S4 and S25 exhibit a more introverted disposition.

Regarding family background, the families of S4, S21, and S25 are supportive and collaborate effectively with the teacher. In contrast, S8's family demonstrates very high expectations, while S18's family tends to deny the presence of behavioral difficulties and holds views that conflict with the teachers.

When students' participation in lessons other than the P4C workshops was examined, all students except S21 were generally willing to take turns speaking and participate in classroom activities. Student S21, however, participated only when feeling entirely confident about the content.

Role of the Researcher/Practitioner and Ethical Considerations

By design, action research is grounded in collaborative processes in which participants work together to identify a problem, implement an action, reflect, and evaluate (Bargal, 2008; Winter, 2005). In the present study, the classroom teacher conducted the action research, serving as both researcher and practitioner. Prior to the initiation of the action research, throughout the implementation process, and following the completion of the study, the researcher/practitioner held meetings with students and parents, conducted the workshops and classroom practices, and organized the checklists, audio recordings, and visual data. Throughout the action research process, the researcher analyzed the collected data, conducted an extensive literature review, and structured the action cycles by seeking support from families and the school guidance counselor.

From an academic perspective, the researcher/classroom teacher has conducted studies using various research methodologies, with the results published in national and international journals. In addition, the researcher has presented action research-based papers at international symposia and participated in workshops related to qualitative research. Professionally, the researcher has 16 years of teaching experience as a primary school teacher and has worked in schools representing diverse socio-economic and socio-cultural contexts. Furthermore, the researcher holds a *Philosophy for Children (P4C) Trainer Training* certification.

Prior to the study, the school administration was informed, and written informed consent was obtained from parents. Throughout the research process, participants' personal information was kept confidential, and a coding system (e.g., S1, S2, S3) was used to ensure anonymity. In addition, ethical approval was obtained from the Social Sciences Ethics Committee of a public university (Approval No. 2024/88).

Data Collection

Data were collected using a checklist, semi-structured interview forms, field notes, and audio and video recordings. In addition, the researcher/classroom teacher maintained a teacher journal throughout the entire research process.

Checklist

Checklists focus on whether a particular behavior or performance occurs and allow comparisons of a student's performance across different time points (Airasian, 1994). In this study, a checklist was used after each P4C workshop to observe students' thinking skills and levels of self-expression. The criteria used to evaluate performance included taking turns speaking in class, developing opposing viewpoints, changing one's ideas during the process (cognitive flexibility), and proposing new ideas.

These criteria were initially drafted using clear and measurable statements. To ensure content validity, the draft checklist was submitted to two field experts for review. Based on their feedback, necessary revisions were made, and the final version of the checklist was developed. Following each implementation, the researcher analyzed the completed checklists to monitor students' developmental progress. A sample checklist is presented in Appendix 1.

Semi-Structured Interview Forms

Interviews, one of the primary data collection tools in qualitative research, can be classified in various ways according to their format, number of participants, or application techniques (Rubin & Rubin, 2012). In the present study, two different interview forms were designed. In terms of format, the interviews were semi-structured; in terms of participants, they were conducted one-to-one; and in terms

of technique, they were conducted face-to-face. The interview forms used with the school guidance counselor and the parents of the participating students are provided in Appendices 2 and 3, respectively.

Field Notes

During the research process, various challenges may arise in the field, or researchers may encounter noteworthy or unexpected situations (Coffey, 1999). Records the researcher keeps of observed, unsystematic data are referred to as field notes (Fetterman, 2010). Although field notes are often understood as written records, digital recordings, visual materials, and audio documents are also considered forms of field notes (Schensul & LeCompte, 2013).

As the researcher was also the classroom teacher, she had the opportunity to observe the class in its natural setting during both instructional and non-instructional times and accordingly kept field notes. In addition, the workshop processes were documented, and data from photographs and video recordings were integrated to provide a more comprehensive dataset.

Teacher Journal

In crowded environments, it may not always be possible to take notes during observations, which may raise questions among participants about why notes are not being taken (Emerson et al., 2008). In some cases, note-taking or recording activities may negatively influence participants' natural behaviors (Bryman, 2012). Under such circumstances, researchers may record participants' conversations, cues, and observations in the form of a journal after each implementation and observation session (Fetterman, 2010).

In the present study, after conducting the P4C workshops, the researcher/classroom teacher regularly organized field notes and audio and video recordings through a teacher journal, drew necessary inferences, and used these data as a basis for planning subsequent phases of the action research.

Action Plan and Implementation

Initially, an action plan was developed to enhance students' critical, creative, caring, and collaborative thinking skills. While designing the action plan, students' grade level was taken into consideration, and appropriate illustrated children's books were selected. Supplementary materials, drama activities, images, and music for use during the warm-up, preparation, and consolidation stages of the P4C workshops were also planned.

Following the first implementation, it was observed that not all students felt comfortable expressing their thoughts, and some hesitated to do so. To clarify whether this situation was independent of the teacher and the workshops or a reaction to them, another teacher (the school guidance counselor) conducted classroom observations and subsequently met with the researcher. Based on field notes, teacher journals, checklists, and an interview with the guidance counselor, it was determined that five students required additional support. At this stage, meetings were held with the parents of the identified students, and the guidance counselor conducted one-on-one sessions with the students.

To foster the development of thinking skills across the entire class and facilitate self-expression among the identified students, a new action plan was developed, and the cyclical process continued in line with the nature of action research (Curtis et al., 2010).

The stages of the P4C process (warm-up, stimulus, community of inquiry, evaluation, etc.) were implemented simultaneously with the whole class. Similarly, drawing activities were conducted collectively; volunteer students took turns speaking, discussed their drawings, and shared their products with their peers. However, during this process, the five identified students were observed in greater detail and were particularly encouraged, through dialogic interactions, to talk about and reflect on their drawings.

Each workshop addressed different themes using different illustrated children's books. Detailed information regarding the P4C workshops conducted throughout the study is presented in Table 2.

Table 2. Illustrated children's books and themes used in the P4C workshops

Implementation	Book Title	Theme
1st Implementation	The Pigeon's Poop	Problem-solving skills, empathy for animals
2nd Implementation	I Am Not Lazy	Diligence, temperament
3rd Implementation	There Is a Wall in the Middle of This Book	Prejudice, cooperation
4th Implementation	How Does It Begin?	Existence, boundaries
5th Implementation	The Black and White Factory	Respect for differences, prejudice

According to Table 2, five different illustrated children's books were used in the study. The themes addressed in these books include prejudice, cooperation, diligence, temperament, philosophy of existence, the boundaries of matter, problem-solving skills, empathy for animals, and respect for differences. In selecting the books, criteria such as the study's purpose, target age group, thematic depth (openness to philosophical inquiry), linguistic and narrative features, and cultural appropriateness were considered.

In the P4C workshops, the stages of warm-up activities, interactive reading, and community of inquiry, including thinking and question-answer activities, lasted an average of 4 class periods.

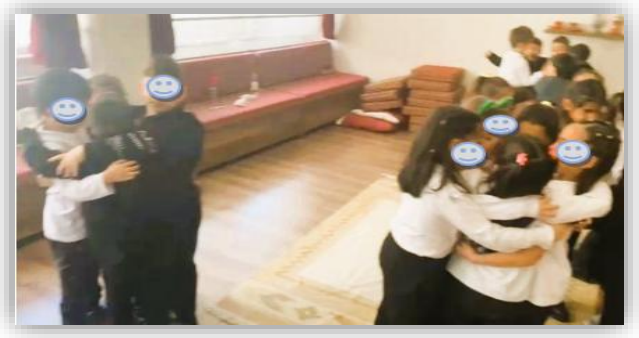


Figure 2. Warm-up activity 1

Figure 2 presents a snapshot from a drama-based warm-up activity focused on cooperation, aligned with the book's theme, *There Is a Wall in the Middle of This Book*.

Figure 3. Warm-Up Activity 2

Figure 3 shows students reenacting Rembrandt's *"The Anatomy Lesson"* as a preparatory activity for the philosophy-of-existence theme addressed in the illustrated children's book *"How Does It Begin?"*

Figure 4. Warm-up activity 3

Figure 4 depicts the dramatization of a Nasreddin Hodja anecdote involving a donkey during the warm-up phase of the P4C workshop, designed to address the theme of empathy for animals associated with the illustrated children's book *The Pigeon's Poop*.

Following the warm-up activities, the illustrated children's book was read aloud as a stimulus. Questions were generated from the text, and a community of inquiry was conducted around them. At specific points, drawings were used either during the book reading or throughout the community of inquiry phase to help students express their thoughts.

Ensuring Validity and Reliability

In qualitative research, the concepts of *credibility* and *transferability* are used as counterparts to validity (Lincoln & Guba, 1985). To enhance credibility, strategies such as data triangulation, prolonged engagement, the inclusion of an external observer for verification purposes, and negative case analysis aimed at improving procedural shortcomings may be employed (Guba, 1981). In the present study, the following steps were taken to ensure validity and reliability.

Triangulation: Triangulation, also referred to in the literature as the use of multiple data sources, is based on the principle of combining different data collection methods within a single study (Silverman, 2013). Accordingly, multiple data sources—including checklists, semi-structured interview forms, field notes, and a teacher journal—were employed in this study.

Prolonged Observation: Prolonged observation enhances the scope validity of qualitative research and allows researchers to identify novel, meaningful, and detailed insights through sustained engagement in the field (Lincoln & Guba, 1985). As the researcher also served as the classroom teacher, opportunities were available to observe students in depth during both instructional and non-instructional contexts.

Peer Debriefing: Peer debriefing refers to the process of sharing research findings with colleagues, field experts, or academics to obtain feedback and ensure objectivity in the research process (Miles & Jozefowicz-Simbeni, 2010). In this study, the school guidance counselor was consulted to identify students for in-depth observation, and the findings were collaboratively compared and evaluated.

Negative Case Analysis: Negative case analysis involves revisiting and revising research procedures in response to unfavorable situations encountered during the study and developing alternative solutions to address inadequacies identified in the data (Erlandson et al., 1993). In the present study, the researcher observed insufficient participation during the initial workshops and responded by developing alternative strategies, ultimately incorporating drawings as a medium for expressing thought.

Member Checking: Member checking is the process by which researchers' analyses are reviewed by participants or individuals closely associated with them, and it is particularly relevant in studies involving young or older participants (Braun & Clarke, 2013). In this study, to verify observations and obtain more detailed insights regarding the five students identified as having difficulty expressing their thoughts, interviews were conducted with their parents.

In qualitative research, reliability is addressed through the concepts of *dependability* and *confirmability* (Lincoln & Guba, 1985). From a qualitative perspective, dependability is achieved when different researchers analyzing the same data reach consistent and coherent conclusions (Franklin et al., 2010). In participatory action research contexts, where the researcher also serves as the classroom teacher and direct comparison across researchers is not feasible, adherence to a step-by-step replication principle is essential. Accordingly, the researcher meticulously documented all stages of the research process, including interviews, implementations, and data analyses, to ensure reliability.

Transferability and credibility represent additional critical considerations for establishing validity and reliability in qualitative research. In line with this approach, data were supported with visual materials, direct quotations from interviews with students, parents, and the school guidance counselor were included, and the findings were reported in rich and detailed descriptions.

Findings

This section presents the findings related to the development of first-grade students' ability to express their thoughts as a result of the P4C workshops implemented in the study.

First Implementation (April 1-12, 2024)

In the first week of the study, the aim was to develop elementary school students' thinking skills, and in line with the action plan, the picture book titled "Güvercin Kakası" (Pigeon Poop) shown in Figure 5 was used.

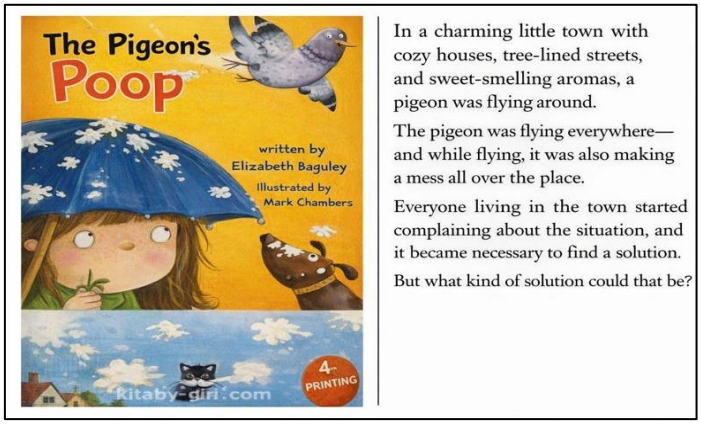


Figure 5. Picture book used in the first implementation

Preparation and relaxation phase: A humorous anecdote emphasizing the bond between Nasreddin Hodja and his donkey was narrated in the classroom. Subsequently, the students were divided into small groups, and each group dramatized the anecdote. This activity aimed to prepare students emotionally and cognitively for the topic.

Community of inquiry phase: The picture book *The Pigeon's Droppings*, which centers on animal compassion, was read using interactive reading strategies, including predicting, questioning, mental imagery, and connecting to prior knowledge and real-life experiences. At the midpoint of the reading process, the following questions were posed to initiate students' thinking:

- What do you think could be done to get rid of the pigeons' droppings?
- If everyone is complaining, is the pigeon to blame?
- Do you think the pigeon is doing this intentionally?

Following the question-and-answer stage, the remaining part of the book was read, and a discussion focused on the solution proposed by the little girl in the story. During this discussion, the theme of animal compassion was further explored. Based on the students' responses, key concepts related to animal compassion were identified, and a concept map was created, as presented in Figure 6.

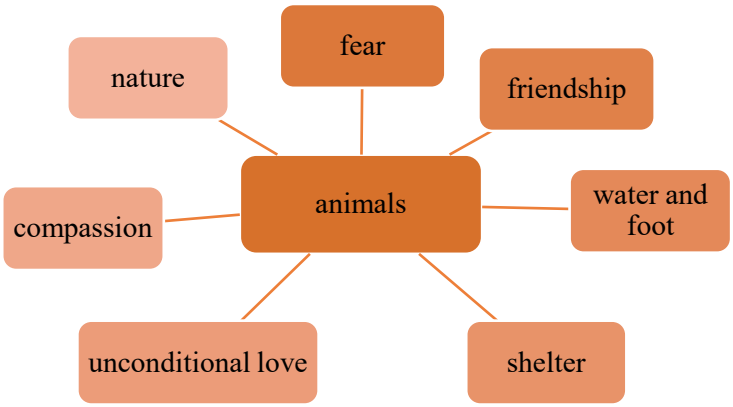


Figure 6. Concept map formed as a result of the first implementation

Results: Following the workshop, observation records, video recordings, and audio recordings were examined, and it was determined that participation across the class was not homogeneous. It was observed that four students consistently wished to participate, nine students participated occasionally, seventeen students raised their hands once, and some students did not express any opinions at all. In addition, based on the data obtained over the one week, the proportions derived from the completed checklists are presented in Figure 7.

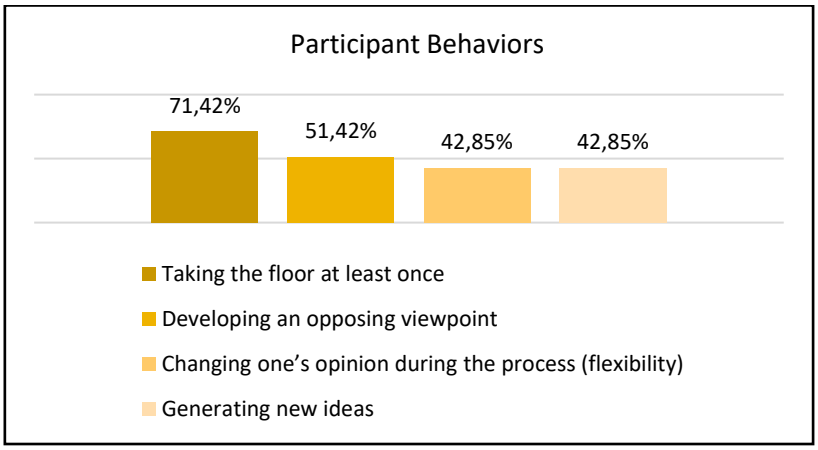


Figure 7. Participant Behaviors in the First Implementation

Figure 7 presents the rates of participants' engagement in the specified behaviors. Of the participants, 71.42% (n = 25) took the floor at least once, and 51.42% (n = 18) demonstrated counter-argument development. The behaviors of changing one's opinion during the process (flexibility) and putting forward a new idea were observed at equal rates (42.85%, n = 15). In addition to the data obtained from the checklist, the following notes from the teacher's journal are presented:

"They listened attentively during the interactive reading phase; however, during the question-answer stage, participation was limited to certain students. Some may genuinely have no ideas, but I think some students are reluctant to express their thoughts." (Classroom teacher, 07.04.2024)

In light of the data obtained, the researcher decided to examine a select group of students more closely to both enhance the thinking skills of all participants in P4C activities and increase the involvement of students who had demonstrated limited participation. After reviewing the checklists and audio and video recordings, the researcher considered that students coded S4, S8, S18, S21, and S25 might be experiencing difficulties expressing themselves. Accordingly, these students were also observed in other lessons to understand the issue better. The teacher's notes include the following:

"Today, we worked on a text in the Turkish lesson. I asked the students to write answers to reading comprehension questions in their notebooks. S18 and S25 were eager to read their answers aloud by raising their hands. However, the other three students did not raise their hands. When I checked their notebooks, one had written an incomplete answer, while the others had answered correctly. Nevertheless, they did not want to speak. When I asked S21 why, the student said that I would not choose them anyway and that it would not be their turn in the classroom." (Classroom teacher, 10.04.2024)

To identify how the five selected students interacted with a different teacher, the researcher sought support from the school guidance counselor, who conducted an activity in the classroom during one lesson and observed the students. The guidance counselor's observations were as follows:

"In general, there are no significant problems in these students' relationships with their peers. S29 was disengaged throughout the lesson and focused on another activity, making it difficult to capture their attention. S18, who is quite active, frequently made remarks to classmates and spoke without permission during the lesson. However, when asked a question related to the activity, the student did not respond—most likely due to fear of being ridiculed. S21 appears to want to raise their hand but does so very subtly. I want to meet with this student individually." (Guidance counselor, 11.04.2024)

Finally, the researcher observed the student mentioned by the guidance counselor and noticed similar patterns in other lessons. As a result, a meeting with the family was planned. With the mother's consent, an audio-recorded interview was conducted at the school. In one part of the interview, the mother stated:

"I have two older daughters. One is in high school, and the other started university this year. Both have had successful academic lives and never caused me any concern. I am striving for my youngest daughter to be the same." (Mother of S21, 12.04.2024)

Based on these and similar statements, it was concluded that the family's expectations of the student were high, which created anxiety in the child. When the child feels that they are not good enough, they hesitate to express their thoughts. As a result, the reasons why the five identified students did not express their thoughts are shown in Figure 8.

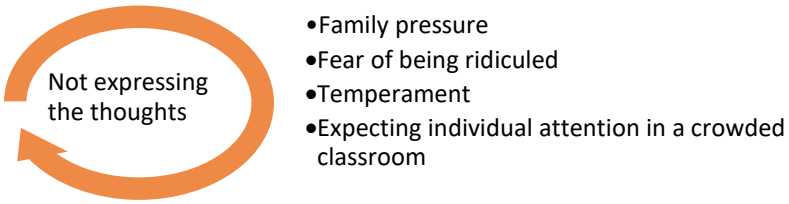


Figure 8. Reasons for Not Expressing Thoughts

When the interviews conducted with the school counselor, parents, and students, together with the data obtained from the researcher's field notes and the checklist, were examined, it was concluded that 5 out of the 35 students in the classroom experienced difficulty expressing their thoughts during the P4C workshops. These difficulties were determined to stem not from physiological problems but from environmental pressure, insufficient opportunities to participate, and individual personality characteristics. In line with these findings, the P4C practices were reviewed and revised for the subsequent implementation.

Second Implementation (15.04.2024-24.04.2024)

To address the problems identified during the first week, the workshop stages shown in Figure 9 were planned.

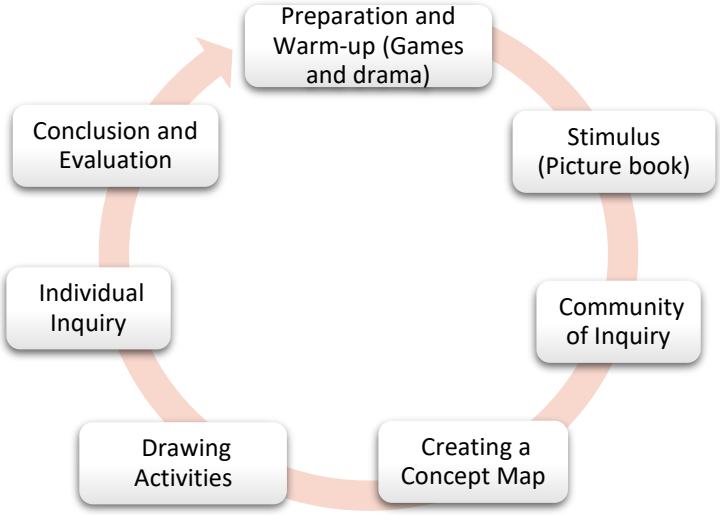


Figure 9. Revised Workshop Process

According to Figure 9, the workshop began with game- and drama-based activities as a warm-up. Subsequently, the selected picture book was read interactively, and the community of inquiry phase was initiated, during which a concept map was created through question-and-answer activities. At this stage, students were additionally asked to produce drawings inspired by the book, and individual interviews were conducted based on these drawings, particularly with students who experienced difficulty expressing their thoughts. Finally, all data were collected, and the workshop was evaluated.

In the second workshop, the picture book titled *"I Am Not Lazy"* was used as the stimulus. Brief information about the book is presented in Figure 10.

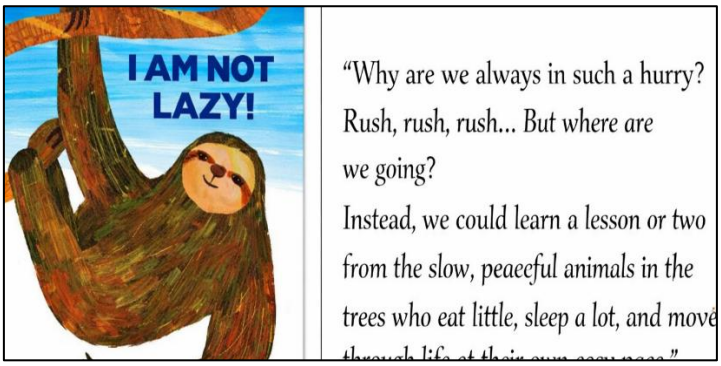


Figure 10. Picture Book Used in the Second Implementation

Preparation and relaxation: During the preparation stage, an ice-breaking activity was conducted through a drama exercise called "statue." Afterwards, the story was read to the students using interactive reading techniques.

Community of inquiry: After observing that the sloth spends the entire day being inactive and that the surrounding animals continuously ask why it is slow, why it speaks slowly, and why it is quiet, the students were asked, *"What do you think laziness is?"* The concept map, formed by writing the students' responses on the board, is shown in Figure 11.

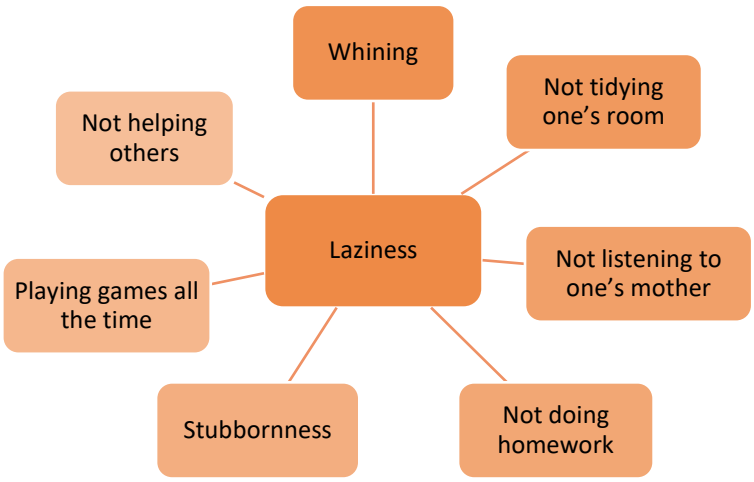


Figure 11. Concept map obtained in the second implementation

Following the continuation of the community of inquiry, the book was completed, and students were asked to describe the sloth's characteristics. At this stage, students used concepts such as *slow*, *speaks little*, and *quiet* when describing the sloth. The inquiry process then continued with the following questions:

- Does doing something quickly and rushing mean being hardworking?
- Is it necessary to always be hardworking?
- Is the sloth sluggish?
- Is the sloth unhappy?
- Do we need to work constantly in order to be happy?
- Does resting make us unsuccessful?

After the inquiries, an attempt was made to develop perspectives on the workshop themes of *hard work* and *temperamental differences*. The community of inquiry phase was concluded with the understanding that some people may behave differently by nature and that this does not mean they are inadequate.

Individual inquiry: It was observed that the five previously identified students again experienced difficulty expressing their thoughts during this workshop, and a new method was implemented as previously planned. Accordingly, all students were asked to draw pictures of what they would say to the sloth if they encountered it. While examining all students' drawings, in-depth interviews were conducted, particularly with the five students who had difficulty expressing their thoughts, and discussions were held based on their drawings. An example of a student's drawing is presented in the figure.

Figure 12. Student drawing from the second implementation (1)

In Figure 12, the student makes a constructive criticism by saying, *"I think you are brilliant, but it would be better if you were not slow,"* and in response, the sloth replies with a happy facial expression, *"You are such a sweet child, I love you very much."* The inquiry then continued as follows:

F (facilitator): Why did you think the sloth was wise?

S4: Because it said that it is happy.

F: Are smart people happy?

S4: They are hardworking; everyone likes them.

F: Do you only like your hardworking friends?

S4: No. Z… has difficulty reading, but she is my closest friend; I love her.

In the continuing dialogue, issues that the student considered when choosing friends were discussed, and it was observed that the student developed a new perspective and was able to express it. Another student in the target group, S8, produced the drawings shown in Figure 13 as a result of the inquiry.

Figure 13. Student drawing from the second implementation (2)

In Figure 13, the child tells the sloth, *"You are very slow, but even this way you do everything very well,"* and the sloth drawn opposite the child is depicted with a happy facial expression. A part of the dialogue between the facilitator and the student is as follows:

F: Why do you think the child wanted to say this to the sloth?

S8: Because everyone calls it lazy, but it is not.

F: Isn't it lazy?

S8: No, it is just slow.

F: Do you think it was upsetting that people did not understand it?

S8: No, it was not upset because the book says, "This is how I am."

F: If it were you, would you be upset?

S8: I do not know. (hesitant silence)

F: For example, how would your friends' behavior make you feel upset?

S8: If they did not include me in the game, if they sulked, or if they laughed at me.

As can be inferred from the student's statements, although hesitation occasionally persisted, S8 showed improvement in expressing thoughts.

Conclusion: After completing the community of inquiry and individual inquiries, the researcher reviewed the collected data. The data regarding the changes in participants' thinking skills after the second implementation are presented in Figure 14.

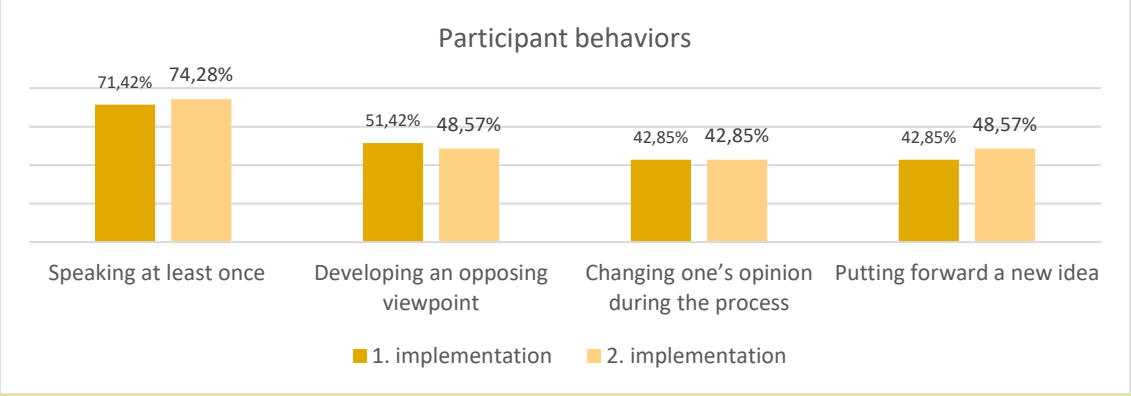


Figure 14. Participant behaviors in the second implementation

The behavior of speaking at least once was observed in 71.42% (26 students) during the first implementation and 74.28% during the second. This rate indicates that the majority of participants actively took part in the discussions.

The behavior of developing an opposing viewpoint remained at a similar level across implementations, occurring at 51.42% (17 students) in the first implementation and 48.57% in the second. This suggests that critical thinking behaviors did not change significantly between the two applications.

The behavior of changing one's opinion during the process remained constant in both implementations at 42.85% (16 students). This indicates that participants were open to revising their ideas to a certain extent, but this openness did not increase.

The behavior of putting forward a new idea was observed at 42.85% (15 students) in the first implementation and increased slightly to 48.57% (17 students) in the second implementation. This finding points to a modest increase in the tendency to make creative contributions.

Based on the checklist data and overall class results, it was determined that three of the five previously identified students participated in dialogue by taking the floor, and two of them generated

new ideas. In particular, Student S4, who has a diagnosed anxiety disorder and minimal interaction with peers, expressing thoughts related to the concept of friendship was considered an indicator of progress. Regarding S4, the following note appears in the teacher's journal:

> *"Since S4 has been raised under strong achievement-oriented pressure from the family, I wanted to engage in a deeper dialogue to explore their thoughts about the concepts of laziness and industriousness. However, I observed that they generally gave short answers and sometimes remained silent, merely nodding. There may be changes in participation across different themes." (Class teacher, 22.04.2024)*

While the teacher expected S4 to participate more actively and generate ideas, given the student's direct connection to the theme, the student responded minimally. The teacher, therefore, decided to continue monitoring the student's development in the subsequent stages of the process.

Third implementation (25.04.2024-05.05.2024)

In the third workshop, a philosophy workshop for children was conducted using the picture book titled "Bu Kitabın Ortasında Duvar Var" (There Is a Wall in the Middle of This Book) shown in Figure 15 as a stimulus.

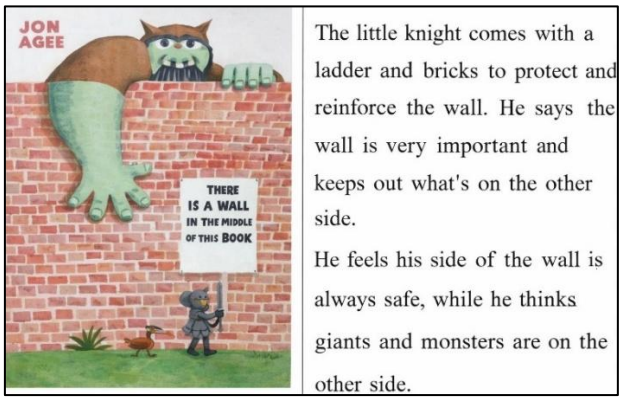


Figure 15. Picture book used in the third implementation

Preparation and relaxation: Based on the central theme of the book, *cooperation*, a drama activity titled "forming molecules from atoms" was carried out during the warm-up phase.

Community of inquiry: After half of the picture book was read, a pause was given and the following questions were posed. Based on the students' responses, the concept map presented in Figure 16 was created.

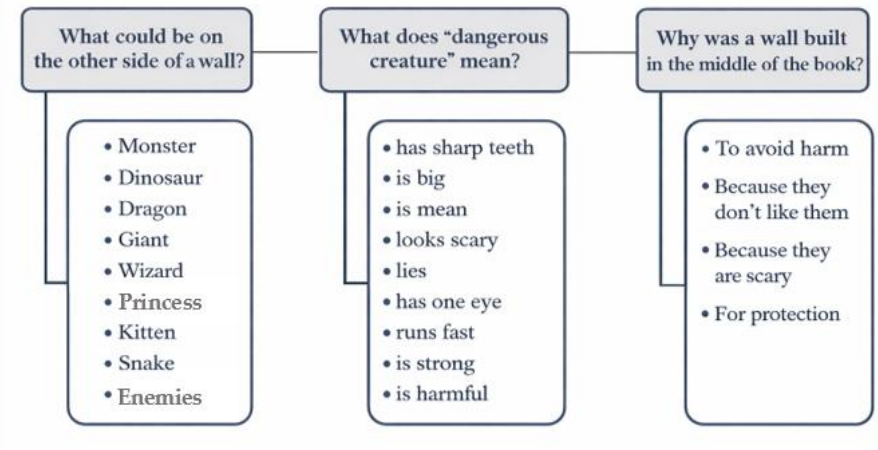


Figure 16. Concept map obtained in the third implementation

According to the concept map, when students were asked to imagine what might be on the other side of the wall, their responses were largely negative and mythical, such as giants, dragons, and wizards. Only two students stated that there might be non-dangerous beings (a kitten and a princess) on the other side of the wall. Views of danger mainly were defined by physical characteristics (teeth, size, eyes), whereas character-based descriptions (e.g., evil-hearted, deceitful) were less frequent. In addition, when students' statements explaining why there should be a wall in the middle of the book were examined, it was found that all emphasized protective reasons. Thus, by the midpoint of the book, the theme of prejudice emerged. At this stage, students were asked to imagine what the other side of the wall might be like.

Individual inquiry: While the community of inquiry began when the book reached its midpoint, it was continued through individual inquiry. After the drawings were completed, individual inquiries were conducted, particularly with the five students who had not actively participated in the inquiry stage. Figure 17 presents a drawing created by student S18.

Figure 17. Student drawing 1 from the third implementation

An excerpt from the dialogue conducted through the drawing with the student coded S18 is as follows:

F (facilitator): Can you describe the picture you drew?

S18: This is a giant.

F: Why are his eyebrows furrowed? Is he angry?

S18: He is evil-hearted and wants to cause harm.

F: How did you understand that he is evil-hearted?

S18: Because he is a giant. Giants are like that.

F: Have you ever seen a giant? How do you know that giants are harmful?

S18: I have not seen one, but I have heard about them.

F: Have you ever heard of a good giant?

S18: No.

The student depicted the giant as large, angry, and evil-hearted based on prior knowledge. After the community inquiry and individual inquiries were completed, the remaining part of the book was read, revealing that the giant on the other side of the wall was actually a good giant who helped and saved the knight, and that the other side of the wall was not as bad as previously assumed. At this stage, the students were asked to draw the giant from the book again on the back of their previous drawings. The drawing belonging to S18 is shown in Figure 18.

Figure 18. Student drawing 2 from the third implementation

A portion of the individual inquiry conducted with the student based on the drawing is as follows:

F: Can you describe what kind of place this is?

S18: On this side of the wall, there are happy animals. Giraffes and monkeys are playing with children.

F: Is there no giant then?

S18: There is a giant, too, but he is a good giant. He took the child out for a walk.

F: How did you understand that he is a good giant?

S18: He saved the child from drowning.

F: If the giant had not been there at that moment and had not seen the child, would he have continued to be a bad giant?

S18: We would not have known him.

F: Are people we do not know bad?

S18: Strangers are bad, but sometimes they can be good too.

In this part of the interview, the student's views on prejudice changed. The student, who did not notice this change or could not express it during the group inquiry, was able to articulate their thoughts more comfortably in the one-on-one interview while discussing their own drawings.

Conclusion: The records related to the community inquiry were examined, and it was determined that participation in the community inquiry increased across the class. In addition, during the reasoning process about what might exist on the other side of the wall, an increase was identified in the rate of developing counter-ideas, such as the possibility of a "princess" or a "kitten," in contrast to peers. Moreover, it was observed that S18 did not exhibit previously observed behaviors, such as laughing at or mocking peers' ideas, during this workshop, and even volunteered to take the floor during the community inquiry stage.

The graph illustrating the development of students' thinking skills after the third implementation is presented in Figure 19.

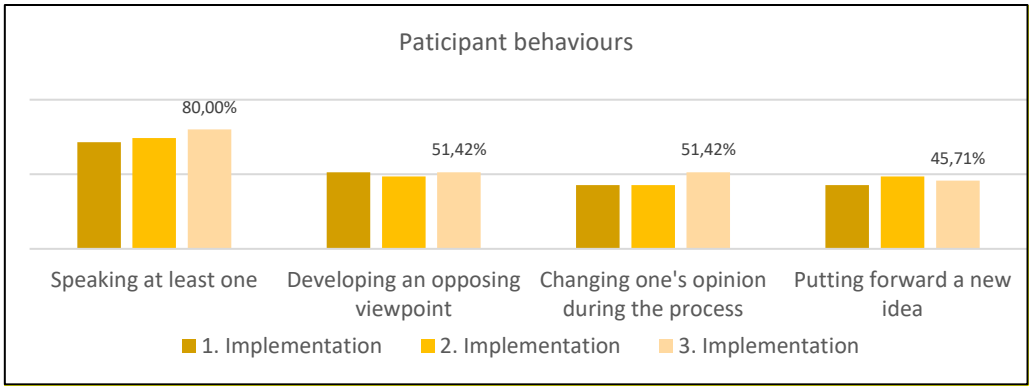


Figure 19. Participant behaviors in the third implementation

An examination of the graph shows that in the third implementation, the greatest increase was observed in the behavior of "taking the floor at least once," while the most notable improvement was in "changing one's opinion during the process." The behavior of "developing counter-arguments" remained relatively stable, whereas the rates of "introducing new ideas" showed slight fluctuations.

Fourth Implementation (06.05.2024-15.05.2024)

In the fourth workshop, the picture book titled *How Does It Begin?* was used. Information related to the book is presented in Figure 20.

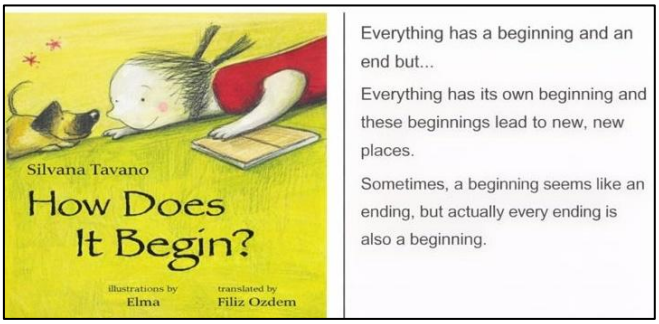


Figure 20. Picture book used in the fourth implementation

Preparation and relaxation: Within the warm-up activities, attention was drawn to human life and living beings through a dramatization of Rembrandt's painting, *The Anatomy Lesson*.

Community inquiry: After the entire book was read interactively, students were asked to provide examples of things that have a beginning and an end. The responses obtained were written on the board, and the concept map shown in Figure 21 was created.

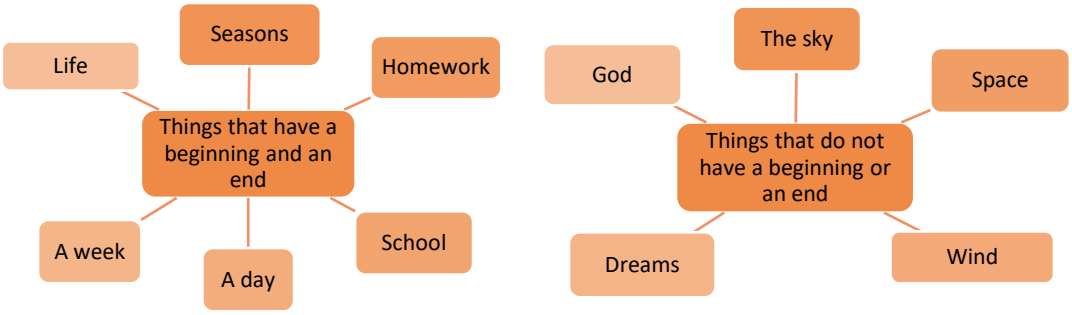


Figure 21. Concept map obtained in the fourth implementation

Subsequently, based on the book stimulus, the following questions were asked:

- Does everything have a beginning?
- Is there something you would like to start?
- Does something that ends begin again?
- What does it mean to have no beginning and no end?

Individual inquiry: After the first stage of the inquiry process was completed, students were asked to draw something that has a beginning and an end. Later, individual interviews were conducted with selected students based on their drawings. Figure 22 presents the drawings of students coded S21 and S4.

Figure 22. Student drawings from the fourth implementation

An excerpt from the dialogue conducted with student S21 is as follows:

F: Would you like to describe your drawing?

S21: There is a baby. First, it is in the mother's womb, then it is born, grows up, gets old, and then dies.

F: Why did you want to draw this?

S21: Because people also have a beginning and an end.

F: In the book, it says that when the seas end, the land begins, and when the land ends, the seas begin. Maybe death is not an end either, but another beginning?

S21: It could be. When my grandfather died, my mother said something like that, too.

In the interview conducted with S21, ontology (the philosophy of being) was discussed through the drawing, based on the book stimulus. In addition, it was observed that the student changed their opinion during the process and moved toward the idea that life is not an end but a beginning for another form of life.

For the second drawing shown in Figure 18, an individual interview was conducted with student S25. A brief excerpt from this interview is as follows:

F: Would you like to describe your drawing?

S25: At the beginning, the world was lush green and ocean blue. However, when the end of the world comes, the greenery will disappear, and environmental pollution will increase.

F: Do you think the end of the world will happen because of environmental pollution?

S25: Yes, because people harm the environment a lot. They pollute the seas and the air.

F: Can something be done to prevent the world from ending like this?

S25: We need to protect the environment and explain this to people.

F: The book says that everything that ends is the beginning of something else. Maybe when the world ends, something else will begin?

S25: But there is no better place for living beings to live.

In the interview with S4, a different perspective emerged: sometimes we need to appreciate what we have, protect nature and the environment, and recognize that the values we lose cannot be replaced.

Conclusion: After the community inquiry and individual inquiries, the researcher analyzed the checklist, video recordings, and audio recordings. The graph obtained from the analysis is shown in Figure 23.

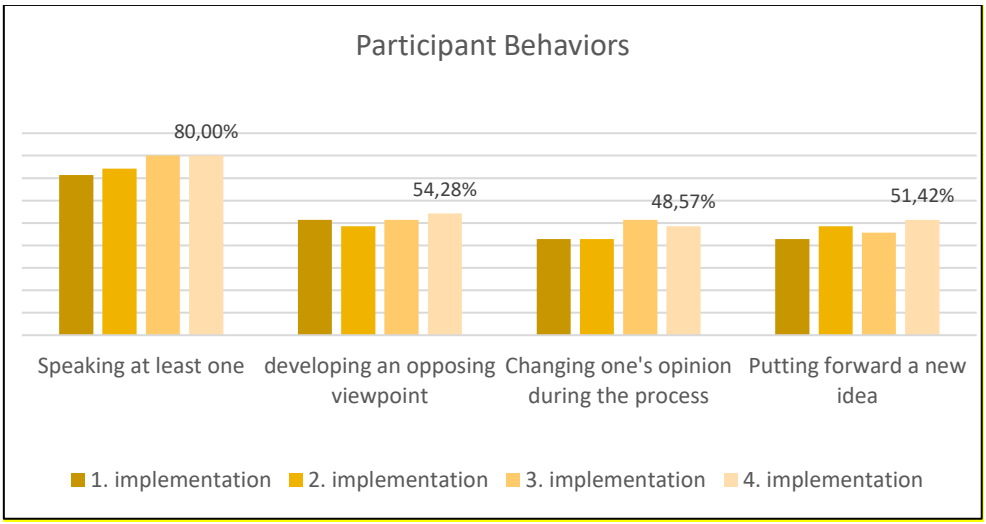


Figure 23. Participant behaviors in the fourth implementation

According to the graph, 80.00% of the participants (n = 28) took the floor at least once. The rate of developing counter-arguments reached the highest level of the study at 54.28% (n = 19). The behavior of changing one's opinion during the process increased compared to previous implementations, reaching 48.57% (n = 17). The rate of introducing new ideas was 51.42% (n = 18), indicating that more than half of the participants made original contributions to the discussion. These findings suggest that participation in the fourth implementation was strengthened both quantitatively and qualitatively.

Regarding the five students who had difficulty expressing their thoughts, an excerpt from the teacher's journal is as follows:

"S21 has always been a student whose thoughts were difficult to identify, as they are usually introverted and quiet. However, when we talked about the drawing they made, I noticed that both the drawing and the explanations were quite in-depth. This process allowed me to conduct a more thorough analysis and to get to know the student more closely."

Fifth Implementation (18.05.2024-29.05.2024)

In the fifth workshop, the picture book titled *The Black and White Factory* was used as a stimulus. An introduction to the book, which is built around the theme that differences, contrary to common belief, enrich our lives and add color, is presented in Figure 24.

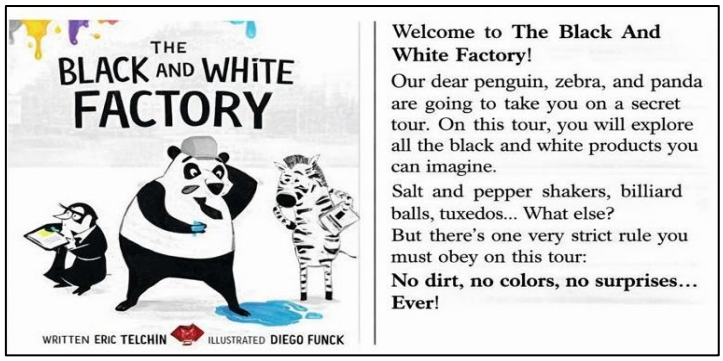


Figure 24. Illustrated children's book used in the fifth implementation.

Preparation and relaxation: A warm-up activity aimed at overcoming a challenge and developing coping skills was conducted through the drama activity called *"Electric Fence."*

Community of inquiry: The entire story was read using interactive read-aloud techniques. As a starting point for the community of inquiry, the question *"What would a world where everything is a single color be like?"* was posed. This was followed by the question *"What would a world with constant change and diversity be like?"* Based on the responses, the concept maps shown in Figure 25 were created, along with the following.

Figure 25. Concept map obtained in the fifth implementation.

As can be inferred from the responses, students generally perceived a monotonous world as ordinary, boring, tiring, frightening, and sad, whereas a diverse world was perceived as fun, adventurous, and easier. Another finding indicated an increase in the rate of generating alternative perspectives within both response groups. In contrast to the dominant view associating a monotonous world with boredom, fatigue, sadness, and ordinariness, some students developed opposing views, suggesting that such a world would be safe.

Similarly, alongside students who believed that a world with diversity would be fun, easy, adventurous, and pleasurable, there was an increase in the number of students who considered that such a world could also be dangerous and tiring and who elaborated on these views through discussion.

The inquiry process was continued with the following questions:

- What would happen if you ate the same food every day?
- What would it be like if all your friends had exactly the same characteristics?
- How would you feel if you suddenly found yourself among people who did not resemble you at all, spoke a completely different language, and whom you had never met before?

Through these questions, students' ways of thinking and their critical thinking skills were aimed to be developed.

Individual inquiry: Following the community of inquiry, all students in the class were asked to choose one color and create a drawing using it. After completing the first drawing, they were then asked to redraw the same picture using any colors they wished. Figure 26 presents the drawings produced by students coded as S4 and S25.

Figure 26. Student drawings from the fifth implementation.

R: Was it more fun to draw using a single color or multiple colors?

S4: Drawing with one color was easier and finished quickly, but I had more fun using many colors.

R: Which one would you prefer in your next drawing?

S4: The multicolored one.

R: Why?

S4: Because I can draw better this way, and it looks more realistic.

The interview with student S4 revealed that the student expressed the view that colors make life more beautiful. In the second drawing presented in Figure 26, part of the interview conducted with the student coded as S25 is as follows:

R: Who is this in the picture?

S25: My friend.

R: Did you like your friend more when they were drawn in one color or when they were multicolored?

S25: Multicolored looks nicer.

R: But the one-color drawing also looks happy. Why might they be happy?

S25: I wanted them to be happy so that they would not feel sad because they are one color.

R: Do you think being one color is something sad?

S25: Yes.

R: If you were made up of only one color and could not be distinguished from your surroundings, what would you do?

S25: I would put a mark on myself so that I could be noticed.

In the individual interview with student S25, the student described monotony as a sad condition and expressed these thoughts comfortably.

Conclusion: When the audio and visual recordings, documents, and checklists related to the philosophical inquiry were examined, it was observed that all five students identified in the community of inquiry dimension actively participated by taking turns to speak. In addition, based on the checklist completed after the final implementation, a graph of the development of all students' thinking skills is presented in Figure 27.

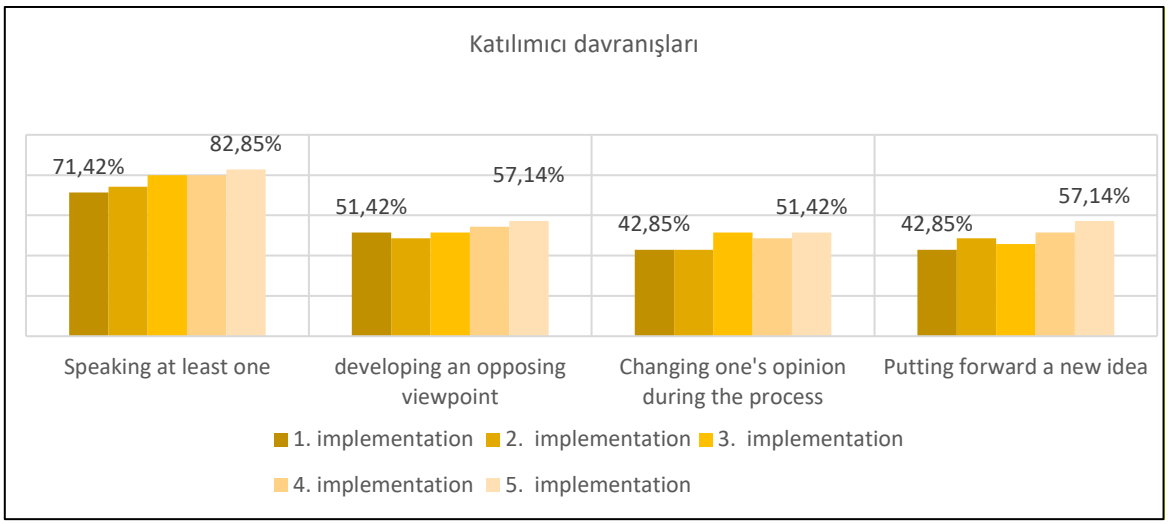


Figure 27. Participant behaviors in the fifth implementation**.**

According to the graph, 82.85% of the participants (n = 29) took the floor at least once during the discussion. The rate of developing opposing viewpoints reached its highest level in the study at 57.14% (n = 20). The behavior of changing one's opinion increased to 51.42% (n = 17) compared to the previous implementations. The rate of introducing new ideas was 57.14% (n = 20), indicating that more than half of the participants made original contributions to the discussion.

When the process from the first to the fifth implementation was examined, it was determined that students demonstrated a clear improvement across all behavioral dimensions.

Discussion

The present study, designed within an action research framework to enhance students' thinking and self-expression skills, was fundamentally structured around Philosophy for Children (P4C) activities. Illustrated children's books were used as stimuli during the workshops, drama techniques were employed in the preparation phase, and drawings were incorporated into the inquiry process to create a multimodal interaction. In this respect, Ferreira (2004) examined the role of stories, dialogue, and activities in teaching basic philosophical skills and found that activities engaging multiple senses also positively influence scientific thinking skills such as observation and inference.

Illustrated children's books were consistently used as stimuli at every stage of the P4C workshops. The first-grade students who constituted the research group had only recently acquired basic literacy skills, and the process was found to significantly enhance their expressive language skills during this critical period, when the foundations of academic achievement are established. The literature also supports the use of children's literature in philosophy workshops, which enriches the thinking environment and enhances student achievement (Çiner & Erginer, 2023; Fisher, 2001; Göksel, 2023; Güvenç & Güney, 2024; Karasu, 2018). In a similar study, Mazı (2008) found that students in the experimental group who participated in thinking activities through stories demonstrated a statistically significant improvement in reading comprehension scores. Likewise, Cassidy and Christie (2013)

reported that philosophical inquiry significantly enhanced children's ability to use meaningful expressions in a study with students aged 5-11. The literature further supports the contribution of thinking education to skills such as providing examples, making connections, and defending ideas (Bülbül Hüner, 2021; Güven, 2019), as well as to the development of reading and writing skills and critical literacy (Pennel, 2012; Ventista, 2019).

Based on the findings obtained from checklists, digital recordings, and field notes, an overall positive development in students' thinking skills was observed at the classroom level. Concept maps were created in line with the themes of the books discussed in each workshop, thereby supporting students' reasoning processes. Gorard et al. (2017) reported that philosophy education for children increases academic achievement and reasoning skills, particularly among students with low academic performance. The effectiveness of P4C activities in terms of comprehension and reasoning, reflective thinking, creativity, and the establishment of relationships between concepts is also supported by the literature (Akkocaoğlu Çayır, 2015; Kefeli & Kara, 2008; Lukey, 2006; Mehdizadeh et al., 2019).

During the research process, particular attention was paid to five students who had difficulty expressing their thoughts. Through verbal dialogues based on students' drawings, efforts were made to facilitate their thinking, and these students showed positive development. Studies in the literature on expressing thoughts through drawings similarly report comparable findings, thus supporting the results of the present study (Atasoy et al., 2007; Murris & Thompson, 2016; Yalçın & Enginer, 2014).

Both the community-of-inquiry and individual-inquiry stages of the study were structured around question-and-answer activities. Analyses revealed increases in both the quantity and quality of responses across each implementation. Similarly, Demirtaş et al. (2018) found that the philosophical inquiry process enhanced preschool students' capacities to ask and answer questions. In terms of dialogue, students who initially preferred to remain silent during the early workshops gradually began to participate actively. In line with this finding, Topping and Trickey (2014) reported an increase in open-ended questions and improvements in classroom dialogue in their study examining the importance of questioning in philosophy for children's workshops.

Another important finding of the study was that P4C activities positively contributed to social relationships and communication skills, both at the classroom level and among the five students who experienced difficulties with self-expression. A student who frequently engaged in teasing toward peers showed a noticeable reduction in this behavior as they began to express their own thoughts more openly. Whitebread et al. (2006) similarly found that philosophy education for children has positive effects not only on cognitive development but also on social skills. Likewise, Gorard et al. (2017) reported that P4C enhances communication skills and supports students' development in respecting different viewpoints, following rules, and solving social problems. The literature also supports the positive impact of philosophy with children's activities for students who experience difficulties in peer relationships, self-regulation, aggressive behaviors, and interpersonal communication (Badri & Vahedi, 2017; Cassidy et al., 2018; Fathi et al., 2020; Jenkins & Lyle, 2010; Jones-Teuben, 2013; Okur, 2008).

Conclusion and Recommendations

This action research study, in which drawings were used to foster thinking in P4C workshops supported by illustrated children's books, yielded the following conclusions.

1. Young students attending the first grade of primary school experience difficulties in expressing their thoughts due to environmental and individual factors, which may prevent P4C activities from fully achieving their intended outcomes. To address this challenge, integrating activities such as drama, music, and drawing into the process helps maintain students' engagement and provides a more comfortable environment for self-expression.
2. Carefully selected illustrated children's books serve as appropriate stimuli for conducting philosophy workshops with young learners. Moreover, applications in which students play with books, dramatize them, and recreate them visually enabled students who had only recently acquired literacy skills to develop positive attitudes toward reading.
3. One-to-one dialogues conducted through drawings revealed that students were able to express themselves more comfortably and that their capacities for asking and answering questions improved. In this respect, the present practice constitutes an effective alternative for supporting the development of expressive language skills.
4. Through the use of drawings as a means of expressing ideas, cognitive activities such as taking turns to speak, generating ideas, developing opposing viewpoints, and openness to new ideas, which were relatively weak in the initial implementation, were observed to increase across the classroom by the final implementation.
5. When students displaying behavioral problems or experiencing communication difficulties were given opportunities for one-to-one interaction and were encouraged to express their thoughts, a reduction in negative behaviors was observed.

Despite these positive outcomes, the study also has certain limitations. While the ideal number of participants in P4C workshops is generally considered to be 15-20, the present study was conducted with a class size of 35 students. To address this limitation, the workshops were not held in a single session but were spread over an extended period.

Based on the findings of the action research, recommendations are presented under three headings.

For researchers: Future studies should examine how the effect of using drawings to support the expression of thinking varies across different grade levels. In addition, longitudinal or follow-up studies may be conducted to assess the sustainability of the gains.

For teachers: Teachers may develop their own strategies to adapt P4C activities for all age groups in order to enhance children's thinking capacities. By carefully analyzing students' needs, teachers can modify and enrich both the process and the content accordingly.

For policymakers: Incorporate thinking education into curricula from an early age, considering its social and cognitive impacts. In line with contemporary educational needs, they should also ensure that in-service training in this field is provided for all classroom teachers.

References

Airasian, P. W. (1994). *Classroom assessment*. McGraw-Hill.

Akkocaoğlu Çayır, N. (2015). *A qualitative study on philosophy education for children* (Doctoral dissertation No. 394850) [Doctoral dissertation, Hacettepe University]. Council of Higher Education National Thesis Center.

Akyol, H. (2006). *Yeni programa uygun Türkçe öğretimi yöntemleri*. Kök Yayıncılık.

Atasoy, B., Kadayıfçı, H., & Akkuş, H. (2007). Öğrencilerin çizimlerinden ve açıklamalarından yaratıcı düşüncelerinin ortaya konulması (Çizimler ve açıklamalar yoluyla yaratıcı düşünceler). *Türk Eğitim Bilimleri Dergisi*, *5*(4), 679-700. https://dergipark.org.tr/tr/pub/tebd/issue/26114/275127

Aydın, A. (1998). *Sınıf yönetimi*. Anı Yayıncılık.

Badri, R., & Vahedi, Z. (2017). The effectiveness of "the philosophy for children program" on the spiritual intelligence of the students. *International Journal of Educational and Psychological Researches, 3*(2), 150-156. https://doi.org/10.4103/2395-2296.204127

Bargal, D. (2008). Action research: A paradigm for achieving social change. *Small Group Research, 39*(1), 17-27. http://doi.org/10.1177/1046496407313407

Baumfield, V., Hall, E., & Wall, K. (2013). *Action research in education*. Sage Publications.

Boyacı, N. P., Karadağ, F., & Gülenç, K. (2018). Çocuklar için felsefe / çocuklarla felsefe: Felsefi metotlar, uygulamalar ve amaçlar. *Kaygı,* (31), 145-173. https://doi.org/10.20981/kaygi.474657

Braun, V., & Clarke, V. (2013). *Successful qualitative research: A practical guide for beginners*. Sage Publications

Bryman, A. (2012). *Social research methods* (4th ed.). Oxford University Press.

Bülbül Hüner, S. (2021). Effect of Socratic questioning on success and attitude and the compliance of the thoughts produced to intellectual standards. *Ankara University Journal of Faculty of Educational Sciences, 54*(1), 35-68. https://doi.org/10.30964/auebfd.632214

Calhoun, E. F. (1994). *How to use action research in the self-renewing school*. Association for Supervision and Curriculum Development. https://eric.ed.gov/?id=ED370205

Cassidy, C., & Christie, D. (2013). Philosophy with children: Talking, thinking, and learning together. *Early Child Development and Care, 183*(8), 1072-1083. https://doi.org/10.1080/03004430.2013.773509

Cassidy, C., Christie, D., Marwick, H., Deeney, L., McLean, G., & Rogers, K. (2018). Fostering citizenship in marginalised children through participation in a community of philosophical inquiry. *Education, Citizenship and Social Justice, 13*(2), 120-132. https://doi.org/10.1177/1746197917700151

Clark, A., & Moss, P. (2011). *Listening to young children: The Mosaic approach* (2nd ed.). National Children's Bureau.

Coffey, A. (1999). *The ethnographic self: Fieldwork and the representation of identity.* Sage Publications.

Collado, F. Y. (1999). *The role of spontaneous drawing in the development of children in early childhood settings* (Report No. ED438898). https://files.eric.ed.gov/fulltext/ED438898.pdf

Corey, S. M. (1949). Action research, fundamental research, and educational practices. *Teachers College Record*, *50*(8), 509-514. https://doi.org/10.1177/016146814905000802

Cox, S., Drew, S., Guillemin, M., Howell, C., Warr, D., & Waycott, J. (2014). *Guidelines for ethical visual research methods*. Melbourne School of Population and Global Health. https://www.semanticscholar.org/paper/Guidelines-for-Ethical-Visual-Research-Methods-Cox-Drew/d146f6da1847dd370b022bd1ee807672b7247249

Curtis, R., Webb-Dempsey, J., & Shambaugh, N. (2010). Understanding your data. In R. P. Pelton (Ed.), *Action research for teacher candidates: Using classroom data to enhance instruction* (pp. 27-57). Rowman & Littlefield Education. https://doi.org/10.5040/9798216408000.ch-2

Çakmak Tolan, Ö., & Genç, M. (2021). Okul öncesi dönemde sık görülen psikolojik ve gelişimsel bozukluklarda oyun terapisi uygulamaları. *Psikiyatride Güncel Yaklaşımlar, 13*(2), 207-231. https://doi.org/10.18863/pgy.757366

Çiçek, H. (2017). Felsefeyle çocuk çocukla felsefe. *Çocuk ve Medeniyet, 2*(4), 51-59. https://dergipark.org.tr/tr/pub/cm/issue/57210/807865

Çiner, H., & Erginer, E. (2023). Çocuklar için felsefe (P4C) çalışmalarında çocuk edebiyatının kullanımı: Sistematik bir derleme. *Kapadokya Eğitim Dergisi, 13*(1), 16-38. https://doi.org/10.37291/kaped.1365387

Çotuksöken, B. (1994). *Felsefi söylem nedir?*. Kabalcı Yayınevi.

Demirtaş, V. Y., Karadağ, F., & Gülenç, K. (2018). Levels of the questions formulated by preschool children during the philosophical inquiry process and the qualities of their answers: Philosophy with children. *International Online Journal of Educational Sciences, 10*(2), 277-294. https://doi.org/10.15345/iojes.2018.02.019

Direk, N. (2019). *Filozof çocuk, çocuklarla felsefe konuşmaları*. Pan Yayıncılık.

Dirican, R., & Deniz, Ü. (2020). Felsefe etkinliklerinin okul öncesi dönem çocukların felsefi tutum ve davranışlarına etkisinin incelenmesi. *Kastamonu Eğitim Dergisi, 28*(1), 421-430. https://doi.org/10.24106/kefdergi.3625

Droit, R. P. (2017). *Çocuklarla felsefe sohbetleri* (A. Karakış, Trans.). Say Yayınları.

Emerson, R., Shaw, L., & Fretz, R. (2008). *Bütün yönleriyle alan çalışması: Etnografik alan notları yazımı* (E. Koca, Trans.). Birleşik Yayınevi.

Eren, M., Erhan, A., & Uluülke, F. (Eds.). (1998). *Alphabetical school encyclopedia* (Vol. 6, pp. 1714–1716). Görsel Yayınları. https://www.nadirkitap.com/alfabetik-okul-ansiklopedisi-cilt-6-gorsel-yayinlar-kitap31680344.html

Erdoğan, P. (2018). *Çocuklarla felsefe yaklaşımının düşünsel, tarihi ve sosyal temelleri üzerine bir inceleme* (Thesis No. 501662) [Master's thesis, Ankara University]. Council of Higher Education National Thesis Center.

Erlandson, D. A., Harris, E. L., Skipper, B. L., & Allen, S. D. (1993). *Doing naturalistic inquiry: A guide to methods.* Sage Publications.

Fathi, L., Ahghar, G., & Naaderi, E. (2020). Teaching philosophy for children (P4C) using a community of inquiry method and impact on students' interpersonal relations. *Quarterly Journal of Education, 35*(4), 51-64. https://qjoe.ir/article-1-2088-en.html

Ferreira, L. B. M. (2004). *The role of a science story, activities, and dialogue modeled on Philosophy for Children in teaching basic science process skills to fifth graders* (Publication No. 3135465) [Doctoral dissertation, Montclair State University]. ProQuest Dissertations and Theses Global.

Fetterman, D. (2010). *Ethnography: Step-by-step*. Sage Publications. https://doi.org/10.4135/9781483348858.n17

Fisher, R. (2001). Philosophy in primary schools: Fostering thinking skills and literacy. *Literacy, 35*(2), 67-73. https://doi.org/10.1111/1467-9345.00164

Franklin, C., Cody, P. A., & Ballan, M. (2010). Reliability and validity in qualitative research. In B. A. Thyer (Ed.), *The handbook of social work research methods* (pp. 273-292). Sage Publications. https://doi.org/10.4135/9781544364902.n19

Gorard, S., Siddiqui, N., & See, B. H. (2017). Can philosophy for children improve primary school attainment? *Journal of Philosophy of Education, 51*(1), 5-22. https://doi.org/10.1111/1467-9752.12227

Göksel, A. (2023). Dijital ortamda düşündürücü etkinliklerle düşünme becerilerini geliştirme. *Korkut Ata Türk Dünyası Uygulama ve Araştırma Merkezi Dergisi, 14*(2), 1276-1291. https://doi.org/10.5152/korkutata.1391205

Guba, E. G. (1981). Criteria for assessing the trustworthiness of naturalistic inquiries. *Educational Communication and Technology Journal, 29*(2), 75-91. https://doi.org/10.1007/BF02766777

Gür, Ç. (2010). Çocuklarla felsefe. *Sosyal Bilimler Dergisi, 8*(2), 43-55. https://dergipark.org.tr/tr/download/article-file/1335717

Güven, B. (2019). *Çocuklarla yapılan felsefe eğitiminde yaratıcı drama yönteminin kullanılması* (Thesis No. 588332) [Master's thesis, Ankara University]. Council of Higher Education National Thesis Center.

Güvenç, E., & Güney, M. (2024). Erken çocukluk döneminde felsefi düşünme ve değerler eğitimi. *Malatya Eğitim Dergisi, 9*(1), 59-78. https://doi.org/10.29228/makusobed.1482578

Haynes, J. (2008). *Children as philosophers*. Routledge. https://doi.org/10.4324/9780203927595

Haywood, H. C. (2020). *Cognitive early education.* Oxford Research Encyclopedia of Education. https://doi.org/10.1093/acrefore/9780190264093.013.971

Hinchey, P. H. (2008). *Action research primer*. Peter Lang Publishing.

Hopperstad, M. H. (2010). Studying meaning in children's drawings. *Journal of Early Childhood Literacy, 10*(4), 430-452. http://doi.org/10.1177/1468798410383251

Jenkins, P., & Lyle, S. (2010). Enacting dialogue: The impact of promoting philosophy for children on the literate thinking of identified poor readers, aged. *Language and Education, 24*(6), 459-472. 10.1080/09500782.2010.495781

Jones-Teuben, H. M. (2013). *I disagree with myself! Creative thinking in a Key Stage 1 community of enquiry* [Doctoral dissertation, Newcastle University]. https://core.ac.uk/outputs/153778669/?utm_source=pdf&utm_medium=banner&utm_campaign=pdf-decoration-v1

Kantarcı, Z. (2013). Sokrates ve eğitim felsefesi. *Mavi Atlas,* (1), 78-89. https://dergipark.org.tr/tr/pub/gumusmaviatlas/issue/7496/98772

Karadağ, F., Yıldız Demirtaş, V., & Yıldız, T. (2017). 5-6 yaş çocuklar için felsefi sorgulama yoluyla eleştirel düşünmenin değerlendirilmesi ölçeğinin geliştirilmesi. *International Online Journal of Educational Sciences, 9(4),* 1025-1037. https://toad.halileksi.net/sites/default/files/pdf/5-6-yas-cocuklar-icin-felsefi-sorgulama-yoluyla-elestirel-dusunmenin-degerlendirilmesi-olcegi-toad.pdf

Karasu, F. Z. (2018). *İlkokul çocuklarıyla felsefe yapmak üzerine nitel bir çalışma: Dördüncü sınıf örneği* (Thesis No. 537665) [Master's thesis, Hacettepe University]. Council of Higher Education National Thesis Center.

Kefeli, İ., & Kara, U. (2008). Philosophical and critical thought development of the child. *Ankara University Journal of Faculty of Educational Sciences (JFES), 41*(1), 339-357. https://doi.org/10.1501/Egifak_0000000193

Kendrick, M., & McKay, R. (2009). Researching literacy with young children's drawings. In M. Narey (Ed.), *Making meaning* (pp. 53-70). Springer. https://doi.org/10.1007/978-0-387-87539-2_4

Lincoln, Y., & Guba, E. (1985). *Naturalistic inquiry: Establishing Trustworthiness.* Beverly Hills. https://doi.org/10.1016/0147-1767(85)90062-8

Lipman, M. (1998). An interview with Matthew Lipman (E. Bosch, Interviewer). *Analytic Teaching*, *18*(1), 1-4. https://journal.viterbo.edu/index.php/at/article/view/666/450

Lipman, M. (2003). *Thinking in education.* Cambridge University Press. https://doi.org/10.1017/CBO9780511840272

Lipman, M., Sharp, A. M., & Oscanyan, F.S. (1980). *Philosophy in the classroom.* Temple University Press.

Lukey, N. L. G. (2006). *Philosophy for Children Hawai'i and its influence on students' reflective thinking in classroom discussions* (Publication No. 545) [Master's thesis, University of Hawai'i]. http://hdl.handle.net/10125/20366

Malchiodi, C. A. (1998). *Çocukların resimlerini anlamak*. Epsilon Yayıncılık.

Matthews, G. B. (2000). *Çocukluk felsefesi* (E. Çakmak, Trans.). Gendaş.

Mavers, D. (2011). *Children's drawing and writing. The remarkable in the unremarkable*. Routledge. https://doi.org/10.4324/9780203844366

Mazı, A. (2008). *Hikayeler yoluyla düşünmenin ilköğretim beşinci sınıf Türkçe dersinde okuduğunu ve dinlediğini anlamaya etkisi* (Thesis No. 215629) [Master's thesis, Mustafa Kemal University]. Council of Higher Education National Thesis Center.

McCall, C. C. (2017). *Transforming thinking: Philosophical inquiry in primary and middle grades* (1st ed., K. Gülenç & N. P. Boyacı, Trans.). Nobel Yayınları.

McKernan, J. (2008). *Curriculum and imagination: Process theory, pedagogy and action research*. Routledge.

Mehdizadeh, A., Araghieh, A., & Heydarieh, H. (2019). The effect of teaching philosophy for children (P4C) on the development of creativity among 5th-grade schoolboy students in the Baharestan education district. *Innovation & Creativity in Human Science, 8*(3), 193-215. https://sanad.iau.ir/en/Journal/ichs/Article/929912

Miles, B. W., & Jozefowicz-Simbeni, D. M. H. (2010). Naturalistic inquiry. In B. A. Thyer (Ed.), *The handbook of social work research methods* (pp. 722-745). Sage Publications. https://doi.org/10.4135/9781544364902.n22

Ministry of National Education. (2018). *İlköğretim programı*. https://mufredat.meb.gov.tr/Programlar.aspx

Ministry of National Education. (2020). *2023 eğitim vizyonu*. http://2023vizyonu.meb.gov.tr/doc/2023_EGITIM_VIZYONU.pdf

Murris, K. (1992). *Teaching philosophy with picture books*. Infonet Publications.

Murris, K. S. (2008). Philosophy with children, the stingray, and the educative value of disequilibrium. *Journal of Philosophy of Education, 42*(3-4), 667-685. https://doi.org/10.1111/j.1467-9752.2008.00640.x

Murris, K. S., & Thompson, R. (2016). Drawings as imaginative expressions of philosophical ideas in a Grade 2 South African literacy classroom. *Reading & Writing, 7*(2), 127. http://doi.org/10.4102/rw.v7i2.127

O'Brien, R. (2001). An overview of the methodological approach of action research. In R. Richardson (Ed.), *Theory and practice of action research*. Universidade Federal da Paraíba.

Oktar, S. Ü. (2019). *Çocuklarla felsefe üzerine bir inceleme*. *Iğdır Üniversitesi Sosyal Bilimler Dergisi*, (17), 45-66. https://dergipark.org.tr/tr/pub/igdirsosbilder/issue/66822/1044951

Okur, M. (2008). *Çocuklar için felsefe eğitim programının altı yaş grubu çocuklarının atılganlık, işbirliği ve kendini kontrol sosyal becerileri üzerindeki etkisi* (Thesis No. 226445) [Master's thesis, Marmara University]. Council of Higher Education National Thesis Center. Malchiodi

Önal, M. (2017). Çocukça felsefe yapmak ya da düşünme ve felsefe dersine hazırlanmak. *Çocuk ve Medeniyet, 2*(4), 61-70. https://dergipark.org.tr/tr/pub/cm/issue/57210/807866

Özdemir, Ö. (2021). *Felsefe sınıfı*. Nobel Yayıncılık.

Özlem, D. (1997). Felsefi hermeneutiğe geçiş yolu olarak tarihselcilik. *Ankara Üniversitesi İlahiyat Fakültesi Dergisi, 40*(2-3), 127-146. https://dergipark.org.tr/tr/pub/auifd/issue/71649/1153439

Pennel, C. (2012). *Epistemic beliefs underpinning discourse within a critical literacy intervention: A multi-case study* [Doctoral dissertation, Cardinal Stritch University]. https://eric.ed.gov/?id=ED552610

Pihlgren, A. S. (2013, May). *Planning for thinking and cognitive development of students* [Paper presentation]. 5th International Conference of Cognitive Science (ICCS), Tehran, Iran.

Republic of Türkiye Ministry of Industry and Technology. (2023). *İlçelerin sosyo-ekonomik gelişmişlik sıralaması araştırması (SEGE-2022)*. Kalkınma Ajansları Genel Müdürlüğü. https://www.kalkinmakutuphanesi.gov.tr/assets/upload/dosyalar/2022-ilce-sege.pdf

Rubin, H. J., & Rubin, I. S. (2012). *Qualitative interviewing: The art of hearing data*. Sage Publications.

Schensul, J. J., & LeCompte, M. D. (2013). *Essential ethnographic methods.* AltaMira Press. https://doi.org/10.5040/9798216409496

Silverman, D. (2013). *Doing qualitative research: A practical handbook* (4th ed.). Sage Publications.

Taş, I., & Dikici Sığırtmaç, A. (2018). Philosophy for children: Fundamental issues and theoretical explanations. In R. Efe, I. Koleva, H. A. Başal, M. Tufan, & E. Atasoy (Eds.), *Educational sciences research in the globalizing world* (pp. 36-52). St. Kliment Ohridski University Press.

Taşdelen, V. (2013). *Felsefeden edebiyata*. Hece Yayınları.

Thomson, P. (2008). *Doing visual research with young children and young people*. Routledge. https://doi.org/10.4324/9780203870525

Topping, K. J., & Trickey, S. (2014). The role of dialogue in philosophy for children. *International Journal of Educational Research,* (63), 69-78. https://doi.org/10.1016/j.ijer.2013.01.002

Turkish Language Association. (2021). *Güncel Türkçe sözlük*. https://sozluk.gov.tr/

UNESCO. (2007). *Philosophy is a school of freedom: Teaching philosophy and learning to philosophize—Status and prospects*. https://unesdoc.unesco.org/ark:/48223/pf0000154173

Taşdelen Wartenberg, T. E. (2018). *Küçük çocuklar için büyük fikirler* (S. Kurtar & A. K. Gülen, Trans.). Sentez Yayıncılık.

Whitebread, D., Linklater, H., Miller, S., Humphreys, K., Glazzard, T., Proudman, S., & Damaskinou, N. (2006). *Teaching thinking: Development of metacognitive abilities in the early years (children aged 3-8) (META)*. University of Cambridge. https://doi.org/10.1007/978-0-387-87539-2_9

Whitfield, P. (2009). The heart of the arts: Fostering young children's ways of knowing. In M. J. Narey (Ed.), *Making meaning: Constructing multimodal perspectives of language, literacy, and learning through arts-based early childhood education* (pp. 153-165). Springer.

Winter, R. (2005). Some principles and procedures for conducting action research. In O. Zuber-Skerritt (Ed.), *New directions in action research* (pp. 9-22). Falmer Press.

Worley, P. (2010). *Felsefe makinesi* (T. Ebesek Büyükuğurlu, Trans.). Paraşüt Kitap.

Yalçın, M., & Erginer, A. (2014). İlköğretim okulu öğrencilerinin okul müdürü algılarına ilişkin yaptıkları çizimler. *Eğitim ve Bilim*, *39*(171), 270-285. https://doi.org/10.15390/ES.2014.1232

Yıldırım, A., & Şimşek, H. (2008). *Sosyal bilimlerde nitel araştırma yöntemleri* (7th ed.). Seçkin Yayıncılık.

Appendices

Appendix 1. Checklist

	Taking the floor at least once	Developing opposing viewpoints	Changing one's opinion during the process	Introducing new ideas	
	Etk.1	**Etk.1**	**Etk.1**	**Etk.1**	**Puan1**
S1	1	0	0	1	2
S2	1	1	0	0	2
S3	1	1	1	1	4
S4	0	0	0	0	0
S5	1	0	1	1	3
S6	1	1	0	1	3
S7	1	1	1	0	3
S8	0	0	0	0	0
S9	1	0	0	0	1
S10	1	1	1	0	3
S11	1	1	1	0	3
S12	1	1	0	0	3
S13	1	1	1	1	4
S14	1	1	1	1	4
S15	1	0	0	0	1
S16	1	0	0	0	1
S17	1	1	1	1	4
S18	0	0	0	0	0
S19	1	1	0	0	2
S20	1	0	1	1	3
S21	0	0	0	0	0
S22	1	0	1	1	3
S23	1	1	1	0	3
S24	1	1	0	1	3
S25	1	1	0	0	1
S26	1	0	1	1	3
S27	1	1	1	0	3
S28	1	1	0	1	3
S29	0	0	0	0	0
S30	1	1	1	0	3
S31	1	1	0	0	3
S32	1	0	1	1	3
S33	1	0	0	0	1
S34 (inclusive education)					
S35 (inclusive education)					

Appendix 2. Parent Interview Form

Based on data from the Philosophy for Children (P4C) workshops, which were conducted as a tool for developing thinking skills, we are conducting this interview with you regarding the student coded as S… Your responses are of great importance to the research we are conducting. This interview will take approximately thirty minutes. All information you provide will be kept confidential, and neither your name nor your child's name will be used anywhere in the study. In order to obtain the data more accurately, an audio recording will be used during the interview. Do you consent to participate in this interview and to the use of audio recording?

Before we begin, if there is anything you would like to ask or add, I am happy to listen.

I am now starting the interview.

1. If I asked you to describe **S…** in a few sentences, how would you describe them?
2. Could you provide information about **S…'s** personality characteristics?
   a. Do you experience disagreements or conflicts of opinion with **S…**? If so, what topics do you most frequently disagree on? How does **S…** express their emotions and thoughts during these conflicts?
   b. In which areas do you feel most in harmony with **S…**? When you reach a joint decision, when **S…** is appreciated, or when they achieve success, how do they express their emotions and thoughts?
3. How would you describe **S…'s** communication with you and with other family members? Could you explain each relationship separately?
4. What are your academic and social expectations for **S…**?
5. Are there any points we have not addressed or that you would like to add?

Appendix 3. Interview Form with the School Counselor

Based on data from the Philosophy for Children (P4C) workshops, which were conducted to develop thinking skills, we are holding this interview with you to gather your views on the class as a whole and on the specific students identified in the study. Your responses are of great importance to the research being conducted. This interview will take approximately thirty minutes. All statements you provide will be kept confidential, and your name will not be used in any part of the study. In order to collect the data more accurately, audio recording will be used during the interview. Do you consent to participate in this interview and to the use of audio recording?

Before we begin, if there is anything you would like to ask or add, I am happy to listen.

I am now starting the interview.

Interview Questions (First Interview)

1. What are your views on the overall classroom atmosphere?
2. What observations have you made regarding students' independent and critical thinking within the classroom?
3. Are there any students whom you think engage in bullying behaviors or are exposed to bullying?
4. Are there students in the classroom who hesitate to express their thoughts?
5. Are there any students whose peer relationships have particularly drawn your attention?
6. Are there any students who have stood out positively or negatively in terms of communicating with you?
7. Is there anything else you would like to add?

Interview Questions (Second Interview)

1. What notable points would you like to share regarding the one-to-one interview you conducted with student coded as S…?
2. In your opinion, is there any situation that would require S… to receive psychological or physical support?
3. Do you think S…'s reluctance to express their thoughts is specific to the P4C workshops, or can it be generalized to other contexts?

Education and Science

Original Article

Cilt 51 (2026) Sayı 226 301-326

The impact of educational digital games on financial literacy in 4th grade social studies courses *

Sinan Özer [1], Ali Ersoy [2]

Abstract

In a rapidly digitizing world, technological developments offer significant opportunities in all areas, including education. The purpose of this study is to examine the effect of educational digital games on the acquisition of financial literacy skills in social studies courses. A mixed research method involving an embedded experimental design was used in this study. The study group consisted of a total of 54 fourth-grade elementary school students attending a public school, selected using criterion sampling and simple random sampling methods. Of these students, 27 were assigned to the experimental group and 27 to the control group. Quantitative data were collected using the Financial Literacy Scale for Elementary School Students and the Financial Literacy Academic Achievement Test, while qualitative data were collected through student journals, field notes, researcher journals, and semi-structured interviews. Because the quantitative data showed a normal distribution, parametric tests were used: t-tests for dependent and independent samples and one-way analysis of variance (ANOVA). Effect sizes were calculated for the results obtained from the paired-sample t-test. Qualitative content analysis was used to analyze the qualitative data. The findings show that teaching through educational digital games increases students' academic achievement, positively affects the sustainability of learning, and improves financial literacy. Furthermore, the qualitative results reveal that students understand the courses better, enjoy themselves during class, enjoy the time they spend in class, and are more motivated when courses are taught using educational digital games. In this study, educational digital games were used in the financial literacy education process. For future research, it is recommended that comprehensive studies be conducted to comparatively evaluate the effectiveness of different teaching methods. Such studies can contribute to determining the most effective teaching strategies for imparting financial literacy skills, especially for educators and policymakers.

Keywords

Elementary school
Social studies
Educational digital games
Digital games
Financial literacy skills

Article Info

Received: 08.08.2024
Accepted: 12.22.2025
Published Online: 04.30.2026

DOI: 10.15390/ES.2026.2512

* This article is derived from Sinan Özer's PhD dissertation entitled "The effect of educational digital games on financial literacy in the 4th grade social studies course", conducted under the supervision of Ali Ersoy.

[1] Ministry of National Education, Trabzon, Türkiye, sinanozer01@gmail.com
[2] Anadolu University, Faculty of Education, Department of Primary Education, Eskişehir, Türkiye, alersoy@anadolu.edu.tr

Introduction

Financial concepts are an integral part of economic life. These concepts are terms that individuals frequently encounter but often struggle to fully understand. These terms can be used in various contexts, both in daily life and in economic discussions. Terms such as borrowing, saving, financial asset management, investment, and income can be given as examples of financial concepts (Öztürk, 2021). Financial literacy refers to the skills necessary for individuals to make informed decisions when faced with increasingly complex financial issues in today's world. It requires knowledge about topics such as credit cards, debt management, investments, and retirement planning, as well as the ability to use this knowledge effectively. Financial literacy is becoming increasingly important and is globally recognized in terms of coping with changes in financial markets and economic difficulties (Adalar, 2019; Organisation for Economic Co-operation and Development [OECD], 2020). Individuals who benefit themselves and the country's economy use financial products correctly, make smart investments, and manage risks effectively. On the other hand, individuals lacking financial knowledge may encounter difficulties and contribute to economic instability. Therefore, conducting risk and return analysis before making investment decisions is crucial (Bayrakdaroğlu & Bilge, 2018). Conducting risk and return analysis also requires the effective use of financial literacy skills.

Today's youth are increasingly forced to take more responsibility for their financial future in an increasingly complex world. Young adults learning to live independently and actively participate in social life must also acquire skills such as budgeting, making sound financial decisions, managing risks, and saving for future uncertainties in their daily lives. Financial literacy is a fundamental life skill for both individuals and societies. It is necessary for making informed financial decisions in daily life and retirement planning. Individuals with high financial literacy skills can make more informed and accurate decisions on financial matters such as money management, bill payments, credit card use, money transfers, budgeting, savings, and retirement savings. Therefore, increasing individuals' level of financial literacy is of great importance for both their personal development and the welfare of societies (Fidancı, 2021).

Financial literacy is recognized as a fundamental life skill for individuals to successfully integrate into society, and financial literacy education plays an extremely important role in helping individuals acquire this skill. Financial literacy education is a process that aims to develop the knowledge, skills, and attitudes necessary for individuals to make informed decisions about financial matters. Government agencies take various steps to protect individuals, include them in the formal financial market, and facilitate their access to this market. One of these steps is to impart financial literacy skills through financial education. Promoting financial literacy has two main objectives: to increase individuals' participation in the financial system and to ensure that they effectively protect themselves when using their financial access. Today, children encounter financial products and services at an earlier age than expected (OECD, 2020). This situation shows that financial literacy can shape children's financial attitudes and behaviors, which can positively affect their financial decision-making skills in adulthood (Zhu et al., 2021). However, it is also emphasized that financial literacy levels among young people are low and that acquiring financial literacy at an early age is important (Amagir et al., 2022; Lusardi, 2019). Without education to increase financial literacy, individuals in society may be unable to use financial products and services adequately due to a lack of financial knowledge. Furthermore, they may not understand the positive contribution these financial products can make to their lives. This situation may lead them to avoid using financial products, negatively affecting the financial market structure of countries (Er & Çetintaş, 2018).

It is stated that the most learning outcomes related to financial literacy belong to social studies, and that this course has the most comprehensive content in this area among compulsory courses (Güvenç, 2017). Social studies education focuses on helping students understand how events unfold and how people relate to each other, rather than just memorizing facts. This education also enables children to discover how to respond to the desires and needs of others. Furthermore, it teaches ways to respect different opinions and cultures. In summary, social studies is a discipline that examines the

political, cultural, economic, and geographical aspects of societies from the past to the present and into the future (Farris, 2004). The social studies course aims to cultivate citizens who have embraced fundamental democratic values, respect human rights, are environmentally conscious, possess critical thinking skills, can make sound decisions, have developed social participation skills, are productive, and are aware of both their rights and responsibilities. One of the most important goals of this course is to cultivate individuals with financial literacy skills. Although financial literacy is a subfield of economics, social studies is one of the most suitable subjects for teaching these skills at the elementary and middle school levels (Arıkan, 2021). Various technology-supported teaching strategies are effectively used in social studies courses. The discipline of social studies is a multifaceted field covering areas such as history, geography, anthropology, archaeology, and politics, and it is of great importance in terms of raising effective citizens with critical thinking and decision-making skills in democratic societies. The use of technology contributes to students' academic skills as well as their awareness of democracy. In this context, technology supports the development of students' higher-order cognitive skills in social studies courses, such as reflective thinking, research, problem solving, critical thinking, and decision making (Baloğlu Uğurlu, 2012; Dai et al., 2022).

Considering the importance of technology use in social studies courses, educational digital games stand out as an effective tool for developing skills within the scope of social studies. Educational digital games aim to develop students' cognitive and affective skills through the use of various technological tools (Aksoy, 2014). Educational digital games can significantly contribute to learning success in education (Coleman & Money, 2019; Mandouit & Hattie, 2023). Some studies show that educational digital games are an effective teaching approach and have positive effects on improving students' academic achievement in various subject areas (Clark et al., 2016). Educational digital games stand out as an important tool in increasing students' learning motivation (Fadda et al., 2022) and in imparting financial literacy skills (Samur, 2016). By providing an interactive learning experience, educational digital games can increase student motivation while also facilitating progress toward specific educational goals.

Consequently, conducting studies in Türkiye to improve students' financial literacy through educational digital games is important in terms of ensuring the effective use of technology in education and imparting financial literacy skills. Educational digital games help students learn abstract financial concepts in a concrete and lasting way (Cheng et al., 2015). These games increase academic achievement while enabling students to be more conscious when making financial decisions (Sabırlı, 2018). It can be said that digital games are particularly effective learning tools for topics such as income, expenses, budgeting, and balance.

Social studies at the elementary school level is one of the most suitable subjects for teaching financial literacy skills. This subject covers historical, social, economic, and political topics, providing students with a multifaceted and holistic perspective. It also facilitates students' connection to the real world through extracurricular activities. Field trips to historical marketplaces, visits to economic zones, and fieldwork conducted with banks contribute to the development of students' financial skills (Akengin & Ersoy, 2015). Teaching financial literacy in social studies courses enables students to acquire skills such as making financial decisions and saving money (Akhan, 2009). It is emphasized that teaching financial literacy as a separate subject in schools or incorporating it into existing curricula provides students with an effective learning opportunity (OECD, 2012). The reason for choosing educational digital games to impart financial literacy skills in this study is the impact of technological developments and digital games on teaching processes. This will help understand the role of technological innovations in equipping students with the targeted knowledge and skills and evaluate the impact of educational digital games on imparting financial literacy skills.

This study, which aims to examine the effect of educational digital games on the acquisition of financial literacy skills in 4th grade social studies courses, seeks to answer the following questions:

1. What is the effect of educational digital game-supported teaching activities on students' academic achievement in financial literacy?
2. What is the effect of educational digital game-supported teaching activities on students' attitudes and behaviors towards financial literacy?
3. What are the views of students on educational digital game-supported teaching activities in terms of gaining financial literacy awareness?

Method

Research Model

This study, which examines the effect of educational digital games on developing financial literacy skills in fourth-grade social studies courses, employed a mixed-methods design. The purpose of mixed-methods research is to combine multiple data sources to provide a deeper understanding of the research problem than studies using a single approach (Guest & Fleming, 2015). Mixed methods are preferred when there is only one data source and the data cannot reflect all aspects of the research, and when more data is needed (Creswell & Plano Clark, 2018). One of the four designs mentioned by Creswell and Plano Clark (2018) for mixed methods research is the embedded experimental design. In this design, the quantitative data collected form the core data set of the research, while the qualitative data are used to support the collected quantitative data set. In some cases, the opposite is true, and quantitative data are used to support qualitative data. The research process is schematically presented in Figure 1.

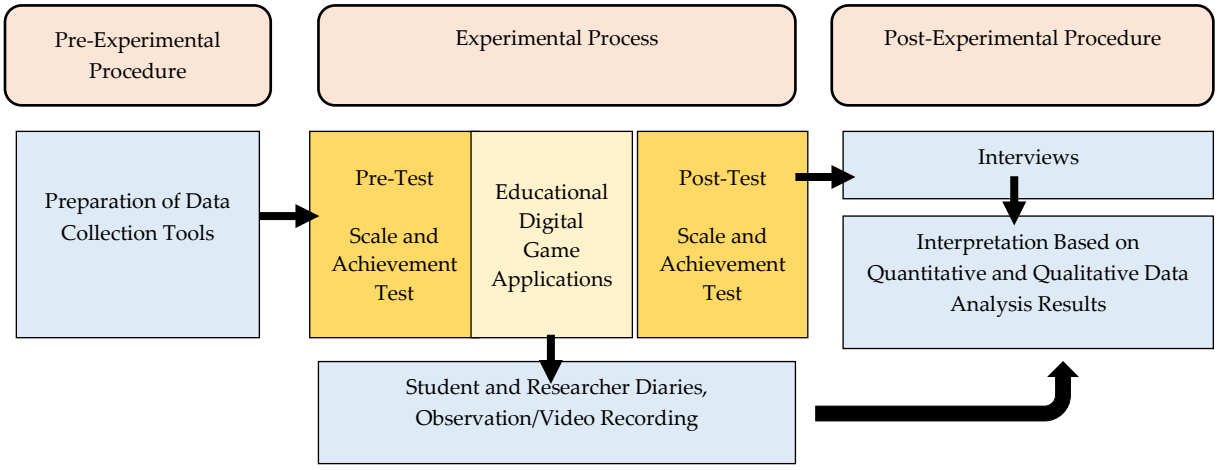


Figure 1. Embedded Experimental Design Research Process

Study Group

Due to the mixed-methods design of the study, different sampling techniques were utilized in the sample selection process. Although financial literacy is an important skill for all age groups, acquiring it at an early age enables individuals to establish a more solid financial foundation in their later lives. Given the importance of acquiring financial literacy skills at an early age, the sample group for the study consists of fourth-grade elementary school students. A criterion sampling technique was used to determine the school where the implementation would be carried out. Yıldırım and Şimşek (2013) stated that the criteria mentioned in criterion sampling can be prepared in advance by the researcher in the form of a list and that all these criteria should be met. The main criteria in this study are the availability of an internet network in the school and the connection speed of this internet network being good when multiple computers are connected at the same time, and the availability of a sufficient number of computers in the school to carry out the a implementation. In accordance with the defined

criteria, a public school located in the district center of A Province was selected for the implementation. To determine the experimental and control group students for the implementation, the Financial Literacy Academic Achievement Test prepared by the researcher (first author) and the Financial Literacy Scale for Primary School Students developed by Özer and Ersoy (2022) were administered as pre-tests to fourth-grade students at the school. According to the results of the tests and scales administered, there was no significant difference between the classes. A lottery was held to determine the experimental and control groups. According to the lottery results, class 4/B (14 girls and 13 boys) was determined as the experimental group, and class 4/D (15 girls and 12 boys) was determined as the control group. The age statistics of the experimental and control group students are given in Table 1. The pre-test results are presented in the findings section.

Table 1. Age Statistics of the Experimental and Control Groups

Grade	AR	$\overline{X}$	S
Experimental Group	9~11	9.81	0.48
Control Group	9–11	9.92	0.47

According to the data presented in Table 1, the age range of the experimental and control group students varies between 9 and 11. The mean age of the experimental group is $\overline{X} = 9.81$, while that of the control group is $\overline{X} = 9.92$.

Data Collection Tools

Quantitative data collection tools

To collect quantitative data in the study, the Financial Literacy Academic Achievement Test and the Financial Literacy Scale for Primary School Students (Özer & Ersoy, 2022) were used from the social studies course "Production, Distribution, and Consumption" learning domain.

Financial Literacy Academic Achievement Test

In order to measure students' success in the study, the Financial Literacy Academic Achievement Test was developed for the social studies course learning domain "Production, Distribution, and Consumption." During the development of this test, the 4th grade Social Studies Course Curriculum was examined to determine the appropriate subject area and learning outcomes for the use and implementation of financial literacy skills within the scope of the study. As a result of the review, it was decided that the learning domain "Production, Distribution, and Consumption," which is directly related to the acquisition of financial literacy skills in the 4th grade Social Studies Course Curriculum, would be selected as the subject domain for the academic achievement test.

Social Studies Course "Production, Distribution, and Consumption" learning domain while preparing the Financial Literacy Academic Achievement Test, the researcher analyzed all the learning outcomes in the unit. These analyses were based on Bloom's Taxonomy, a descriptor table was prepared, and expert opinions were sought. After the learning outcomes were determined, five core learning outcomes were selected, and a total of 40 multiple-choice questions with four options were drafted, with eight questions for each learning outcome. These learning outcomes are: "Makes conscious choices by distinguishing between wishes and needs, recognizes basic economic activities within their family and immediate environment, exhibits conscious consumer behavior as a responsible individual, creates their own budget, and uses resources in their environment without waste." In preparing the question pool, 4th grade social studies textbooks, the Education Information Network (EIN), and other supporting resources were utilized. The draft questions were reviewed by a group consisting of social studies education experts, assessment and evaluation experts, curriculum development experts, and classroom teachers, and feedback was obtained from experts regarding the scope, clarity, and level appropriateness of the questions. Based on the feedback from the experts, distractors in four questions were corrected, and the root of one question was changed. With these adjustments, the 40 draft questions were finalized as. A pilot implementation was conducted with 31 students in the 4th grade to determine the time required for students to answer the test and its level of comprehensibility. Before

the implementation, students were asked to indicate any parts they did not understand, and after the implementation, they were asked to provide feedback on the adjustments that should be made to the test. Students were given 30 minutes for the pilot test (the course duration was set at 30 minutes during the pandemic), and the fastest student completed the test in 21 minutes, while the slowest student completed it in 35 minutes (the second-to-last student completed it in 30 minutes). No intervention was made until the students completed the test, and most students completed the test within 30 minutes: This shows that the time given was sufficient. Feedback from students revealed that there were no problems in understanding the questions and options in the test. In light of these data, it was decided to conduct a reliability study of the draft achievement test. The draft achievement test was administered to 200 students (girls: 105, 52.5%; boys: 95, 47.5%) in the 4th grade of three elementary schools located in the center of District A, Province A, during the 2021-2022 academic year. The test was scored as "1 (one)" for correct answers and "0 (zero)" for incorrect answers. After the implementation, the students' answers were analyzed using the Test Analysis Program (TAP). As a result of the analysis, the students' scores were ranked from highest to lowest, and the top and bottom 27% groups were determined. In the analysis, the discriminative value ("r") and item difficulty indices ("p") of the students' answers to each question were calculated. These analysis results are presented in Table 2.

Table 2. Financial Literacy Academic Achievement Test Item Analysis

Question Number	Difficulty (p)	Item Discrimination (r)	Question Number	Difficulty (p)	Item Discrimination (r)
1	**0.25***	**0.06***	**21**	0.79	0.56
2	0.69	0.46	**22**	0.67	0.68
3	**0.82***	**0.24***	**23**	0.51	0.79
4	0.79	0.45	**24**	0.72	0.64
5	0.72	0.66	**25**	0.72	0.58
6	**0.87***	**0.29***	**26**	0.55	0.68
7	0.77	0.47	**27**	0.54	0.73
8	0.77*	**0.24***	**28**	0.44	0.35
9	0.53	0.62	**29**	0.64	0.64
10	0.57	0.75	**30**	0.80	0.49
11	0.60	0.75	**31**	0.69	0.82
12	0.67	0.62	**32**	0.66	0.58
13	0.52	0.61	**33**	0.60	0.62
14	0.57	0.75	**34**	0.78	0.53
15	0.64	0.42	**35**	0.69	0.68
16	0.71	0.57	**36**	0.71	0.65
17	0.61	0.76	**37**	0.67	0.57
18	0.68	0.48	**38**	0.58	0.80
19	0.79	0.56	**39**	0.74	0.64
20	0.72	0.51	**40**	0.62	0.69
Total				**0.66**	**0.57**

*Questions marked with an asterisk have been removed from the test.

According to the literature, items with a discriminant index below 0.20 are considered to have very low discriminant validity and should be excluded from the test. Items with indices between 0.20 and 0.29 can be revised and reused in the test, while items with indices of 0.30 and above are considered good and can be used without modification (Baykul, 2015). Looking at the discrimination index of the items in Table 2, the index value of question 1 is less than 0.20, while questions 3, 6, and 8 are between 0.20 and 0.29. According to the literature, these questions can be used after correction. Since there were enough questions in the test measuring the same achievement, these four questions were removed from the test. As a result of this adjustment, 36 questions remained. According to Table 2, the item

discrimination indices range from 0.35 to 0.82, and the test's average discrimination index (r) is 0.57, indicating that the test has a good level of discrimination. The item difficulty indices range from 0.44 to 0.80, and the average difficulty index is 0.66. The difficulty index varies between 0 and 1, with questions becoming easier as the index approaches 1 and more difficult as it approaches 0 (Gömleksiz & Erkan, 2010). The reliability coefficient (KR-20) of the test was found to be 0.92, indicating that the test has high reliability (Büyüköztürk, 2014).

Financial Literacy Scale for Elementary School Students

The exploratory factor analysis (EFA) of the Financial Literacy Scale for Elementary School Students (Özer & Ersoy, 2022), developed to determine the financial literacy attitudes and behaviors of elementary school students, was conducted with 294 fourth-grade students attending public schools in District A of Province A. As a result of the reliability analysis, the overall Cronbach's alpha coefficient of the scale was calculated as 0.77. The Cronbach's alpha coefficients for the subscales of the scale ranged from 0.65 to 0.67. In this context, considering that the acceptable lower limit for reliability is determined as 0.60 according to Özdamar (2016), both the scale as a whole and its sub-dimensions can be considered sufficient in terms of reliability. As a result of the EFA, 40.9% of the total variance was explained, and a 3-subdimension scale consisting of 17 items was obtained: *Planned Individual*, *Thrifty Individual*, and *Wasteful Individual.* The scale is a 3-point Likert type, and the score for each question ranges from 1 to 3. The *Planned Individual* sub-dimension consists of 5 items, the *Thrifty Individual* sub-dimension consists of 4 items, and the *Wasteful Individual* sub-dimension consists of 8 items. The score that can be obtained from the *Planned Individual* subscale ranges from 5 to 15, from the *Thrifty Individual* subscale from 4 to 12, and from the *Wasteful Individual* subscale from 8 to 24. High scores reflect individuals' more planned, thrifty, and wasteful behaviors. Sample items are as follows: "I prepare a weekly plan when making my expenses" (*Planned Individual)*, "I save the pocket money I collect during holidays or special occasions" (*Thrifty Individual)*, "I want to buy a product I want even if it is expensive" (*Wastefult Individual)*. Possible score ranges for the *Planned individual* sub-dimension are: 5-7 Points (Low level of planned individual), 8-11 Points (Medium level of planned individual), 12-15 Points (High level of planned individual): Students in this score range are individuals who plan their spending regularly and in detail. For the *thrifty individual* subscale: 4-5 points (low level of thriftiness), 6-8 points (moderate level of thriftiness), 9-12 points (high level of thriftiness): Students in this score range place great importance on saving and carefully managing their financial resources. For the *wasteful individual* subscale: 8-12 points (Low level of wastefulness), 13-18 points (Moderate level of wastefulness), 19-24 points (High level of wastefulness): Students in this score range are individuals who cannot control their spending and generally consume wastefully. After EFA, confirmatory factor analysis (CFA) was performed on 273 fourth-grade students attending public schools in District A of Province A. The χ^2/df value was found to be 1.38 (Ideal Fit: <=2.00, Good Fit: 2.00-5.00), the RMSEA value was 0.023 (Ideal Fit: 0-0.05, Good Fit: 0.05-0.08), the SRMR value was 0.049, and the CFI value was 0.98 (Ideal Fit: 1.00, Good Fit: 0.95-1.00), and TLI value 0.97 (Ideal Fit: 1.00, Good Fit: 0.90-1.00). These values indicate that the model shows ideal fit and provides high fit (Özdamar, 2016). The item factor load values obtained from EFA ranged from 0.439 to 0.840, while the factor load values obtained from CFA analysis ranged from 0.31 to 0.82.

Qualitative data collection tools

In this study, the following qualitative data collection tools were utilized: semi-structured interviews, observations, video and audio recordings, student journals, and researcher journals. These tools were employed to determine students' financial literacy skills, the impact of educational digital games on the development of these skills, students' interactions with the games, and the role of games in the learning process. When preparing the semi-structured interview questions, the literature was consulted, taking into account the purpose of the research. The questions were designed to gain an in-depth understanding of students' financial literacy skill development processes, the impact of educational digital games on these processes, and their knowledge and thoughts. Therefore, sub-questions were also included to ensure the questions were fully understood and to obtain more detailed data. The five questions prepared include experience, knowledge, and opinion questions. For example, the question "Which game(s)/application(s)/activity(ies) did you like the most in financial literacy

courses? Can you give an example? Why?" aims to understand students' experiences of learning financial literacy through games. The draft questions were submitted to experts (two experts in the field of primary education and two classroom teachers) for their opinion. Based on the feedback received from the experts, the necessary adjustments were made to the questions, which were then used in student interviews. In addition, observations, video and audio recordings, student journals, and researcher journals were used as qualitative data collection tools. Observations were made to understand students' interactions with educational digital games, their behavior while playing games, their interest in the games, and their level of focus. Audio recordings were used to record student interviews, ensuring that student statements could be accurately analyzed. Video recordings were used to provide more detail about the observations and to support the interview analysis. Student jounals allowed students to express in writing their experiences with the games, the financial concepts they learned, and how the games helped their learning by giving them a form containing three questions. The researcher's journal was used to record the researcher's observations, analyses, and comments throughout the process, with detailed notes taken on the interviews conducted with students and the behaviors observed. These tools enabled an in-depth analysis of the qualitative data.

Process and Implementation

Course plans for the lessons to be taught were prepared prior to the implementation. The necessary permissions were obtained from the Anadolu University Social and Human Sciences Scientific Research and Publication Ethics Committee prior to the implementation. Prior to the implementation, a pilot study was conducted with 28 elementary school students in the 4th grade at an elementary school located in the center of A District, A Province. The pilot study was conducted over a period of four weeks after obtaining the necessary permissions from the school administration and the classroom teacher. The study was conducted using 8 computers in the school's computer lab. When planning the implementation, an environment reflecting the school environment was created, and possible problems were identified in advance. The games were designed to last an average of three to four minutes each, and a course plan was prepared so that three students would play the games in turn on each computer. No problems related to the internet connection or computers were encountered during the pilot implementation, but high data usage was observed during the implementation. To address this issue, measures were taken to check the internet speed at the school where the experimental implementation would take place and to ensure there were no problems with data usage. A temporary solution using a mobile internet connection was planned in case of any negative circumstances. While teaching the course with educational games and at the end of the course, students were asked for their opinions and feedback was recorded. The feedback received indicated that the time was insufficient in two games, objects falling from above were too fast in one game, the instructions at the bottom of the screen were difficult to read in another game, and one game was very difficult to play. Necessary adjustments were made based on this feedback, and the games were finalized for the experimental implementation.

Digital Game-Supported Instruction Program

The educational digital games used in this study were designed in accordance with the analysis, design, development-implementation, and evaluation stages of the "Spiral Educational Game Design Model" developed by Akgün et al. (2011). During the educational analysis phase, the financial literacy-related learning outcomes included in the 4th grade Social Studies Course Curriculum were identified. Based on a literature review, the financial literacy knowledge that 4th grade students should possess, the purpose of the game, and its content were determined. Accordingly, the content to be used in the game was created. When moving on to the game analysis phase, considering the data obtained from the educational analysis and the limited digital gaming experience of elementary school students, as indicated by, a two-dimensional, individually playable game style was preferred. During the educational design phase, decisions were made on how components such as motivation, interaction, narrative context, motivation, and adaptation would be incorporated into the game. The design process was carried out in collaboration with three social studies education experts and two computer and instructional technology education experts. After deciding on the game design based on the experts'

opinions, the platform on which the game would be developed was determined. In this research, Web 2.0 tools such as Wordwall and Scratch were used in the development of digital games; activities such as matching, object finding, and Jigsaw puzzles were designed on Wordwall, while Scratch was used to create original game content through block-based coding. In addition, ready-made recycling-themed digital games available on the Cevkococuk.org website were also used. The games are available on these sites under the name Production, Distribution, Consumption unit for 4th grade. The stage designs for the games designed during the development phase were created. The developed games were subjected to internal evaluation by taking into account the opinions of primary education, social studies education, and classroom teacher experts. In line with the opinions, adjustments were made regarding issues such as the appropriateness of the visuals and objects used in the game to the level and achievements. As explained above, in the implementation phase, a pilot implementation study of the games was conducted, and the researcher observed how the students played the games and identified any overlooked or errors found by the students. In the evaluation phase, it was decided to use the educational digital games developed based on expert opinions and pilot implementation results in the implementation phase of the research, as the game content was found to be appropriate for the learning outcomes and student level. After the pilot study, the necessary changes were made to the educational digital games, and the implementation process began.

The implementation lasted six weeks. The researcher conducted the educational digital game-supported teaching implementation prepared to impart financial literacy skills to the 4th graders in the experimental group. The courses were conducted in accordance with the prepared lesson plans, with three lessons per week for a total of 18 lessons over six weeks. The last three lessons were conducted for evaluation purposes. The implementation was planned so that two games would be played in each lesson. Which student would play which game on which computer was indicated by a compass hung on the front of the computers. Before playing the games, each game was introduced to the students via the interactive whiteboard. Eight students played the games at each stage. The remaining students waited their turn in their seats. The games were played in a competitive environment, and at the end of the course, small prizes were given to the winners of each game. This motivated the students who were waiting for their turn to play. The students did not find it difficult to play the games in class. Only a few students had problems using the mouse. The researcher took on the role of a guide in the course and intervened when necessary. The setup of the computers in the classroom and the seating arrangements of the students while playing the games are shown in Figure 2. In the control group, courses were conducted in accordance with the existing curriculum for six weeks. The control group followed their courses with their own teachers. Every week, regular meetings were held with the control group's teacher to gather information about how the courses were conducted and the teaching methods used. Courses in the control group were conducted by the teacher using teaching methods that did not include educational digital games. During this process, student participation in the course was ensured through strategies such as lesson narration, question-and-answer activities, and the use of interactive whiteboards. In addition, visual and auditory materials were presented to students in relation to the content of the courses, with the aim of contributing to the understanding of the courses. Visuals from the educational digital games used in the study (Figure 3 and Figure 4) are presented below.

Figure 2. Students' Seating Arrangements at Computers

Figure 3. I'm Getting My Needs Game

Figure 4. Recycling Game

Measures Taken Regarding Validity and Reliability

Since a mixed method was used in the research, both quantitative and qualitative data were utilized. The validity and reliability data for the quantitative data collection tools, namely the Financial Literacy Academic Achievement Test and the Financial Literacy Scale for Primary School Students (Özer & Ersoy, 2022), are presented above under the heading of quantitative data collection tools. For qualitative data, credibility, transferability, consistency, and confirmability studies were conducted (Erlandson et al., 1993). *Credibility:* To ensure credibility in the study, various data collection tools such as pre-implementation, during-implementation, and post-implementation achievement tests, scales, observations, interviews, student journals, and researcher journal were used. This method ensured that data was collected from different sources and that multiple types of data were included in the process. Since the researcher was also the implementer, he spent a long time in the research environment, thus obtaining contextually rich data and conducting comprehensive observations. A systematic approach was adopted in the stages of preparing data collection tools, analyzing data, and finalizing themes. To prevent data loss, courses were recorded on video and audio, and computer screens were also recorded using screen recording software. To increase the credibility of the research, sample excerpts from classroom and screen images are presented in Figure 5.

Figure 5. Images from the Implementation

Transferability: The research process has been described in a detailed and transparent manner. Code names were used to protect the privacy of participants, and direct quotations were presented using these code names. All stages of the research, including details about the pre-research, implementation, and post-research phases, were conveyed with an unbiased approach. *Consistency:* The data obtained from the data collection tools used in the research have been structured in a way that supports each other, and content consistency has been ensured between data sources. Visual and auditory data have been carefully analyzed for accuracy; in particular, video recordings have been watched by the practitioner to confirm the accuracy of the analysis process and prevent possible data loss. Thus, methodological consistency was ensured among the data obtained during the research process, increasing the reliability of the findings. *Verifiability:* In the research, which used both qualitative and quantitative data collection tools, the findings obtained from different data sources were evaluated through comparative analysis. In this context, the existence of meaningful relationships between the themes obtained from qualitative data (interviews, observations, journals, etc.) and quantitative findings (scale and achievement test results) was examined; overlaps, similarities, and differences between the findings were systematically analyzed.

Data Analysis

Both quantitative and qualitative data were collected during the data collection phase of the research. Therefore, both quantitative and qualitative analysis techniques were used in the analysis of the data.

Analysis of quantitative data

The Statistical Package for the Social Sciences 2024 (SPSS) program was used to analyze quantitative data in the study. To determine the tests to be used in analyzing the quantitative data obtained from the pre-test, post-test, and retention test conducted before and after the implementation, the normality assumption of the data was first tested. At this stage, the kurtosis, skewness coefficients, and standard errors of the data groups were examined. In addition, the Shapiro Wilk hypothesis test was used to examine the data. According to Büyüköztürk (2014), the Shapiro Wilk test can be used to test hypotheses indicating whether data obtained from sample groups smaller than 50 are normally distributed. The pre-test, post-test, and retention test mean scores for the Social Studies course "Production, Distribution, and Consumption" learning domain Financial Literacy Academic Achievement Test and Financial Literacy Scale for Elementary School Students (Özer & Ersoy, 2022) show a normal distribution. Due to the normal distribution of the collected data, the parametric dependent and independent sample t-test and one-way ANOVA test were used to compare the experimental and control group data.

Analysis of qualitative data

In this study, qualitative data were analyzed using inductive content analysis. The purpose of content analysis, which can be performed using inductive and deductive methods, is to discover patterns, themes, and categories within the data (Patton, 2014). The video recordings collected during the research process, the interviews conducted with the students, and the data obtained from the researcher and student journals were analyzed using inductive content analysis. After all qualitative data were collected, the open coding process was initiated, and the data were examined repeatedly. Data obtained from student interviews and student journals were analyzed as primary data. Audio recordings obtained from student interviews were converted into written texts and transferred to a digital environment. Subsequently, qualitative data obtained from student journals and interview transcripts were systematically coded. Following this process, the researcher's journals were also analyzed, leading to the identification of categories and themes from the coding. To ensure the validity of the identified themes and the holistic evaluation of the data, video recordings of the implementation process were examined in detail and used as a source to support the analysis findings. The analysis resulted in the identification of the theme "Educational Digital Game-Based Learning Experience" and 6 sub-themes related to this theme. Findings related to the acquisition of financial literacy skills were grouped under the theme "Financial Literacy Awareness," and 9 sub-themes were identified within this main theme.

Findings

Financial Literacy Academic Achievement Test and Financial Literacy Scale for Elementary School Students One-Way ANOVA Test Pre-test Results

The Financial Literacy Academic Achievement Test prepared by the researcher and the Financial Literacy Scale for Primary School Students developed by Özer and Ersoy (2022) were administered as pre-tests. The pre-test results are presented in Table 3 and Table 4.

Table 3. Social Studies Course "Production, Distribution, and Consumption" Learning Domain Financial Literacy Academic Achievement Test One-Way ANOVA Test Pre-Test Results

Grade	n	$\overline{X}$	S	Source of Variance	Sum of Squares	Mean Squares	F	p
4/A Branch	27	19.78	6.84	Inter-group	369,435	123,145	2,301	.082
4/B Branch (E)*	27	18.93	7.58	Within Group	5,565.481	53,514		
4/C Branch	27	20.44	8.00	Total	5,934.917			
4/D Branch (C)*	27	15.63	6.74					
Total	108	18.69	7.44					

*(E: Experimental Group, C: Control Group)

Table 4. One-Way ANOVA Test Pre-Test Results for the Financial Literacy Scale for Elementary School Students

Grade	n	$\overline{X}$	S	Source of Variation	Sum of Squares	Mean Square	F	p
4/A Branch	27	43.70	4.49	Inter-group	73,667	24,556	1,290	.282
4/B Branch (E)*	27	42	3.93	Within Group	1980.296	19,041		
4/C Branch	27	43.89	5.13	Total	2,053,963			
4/D Branch (C)*	27	42.33	3.75					
Total	108	42.98	4.38					

*(E: Experimental Group, C: Control Group)

According to the pre-test analysis results presented in Table 3 and Table 4, it is seen that the responses of all classes to the academic achievement tests and scales are equivalent in terms of academic knowledge ($p > .05$). A lottery was held among these four branches to determine the experimental and control groups. According to the lottery results, class 4/B (14 girls, 13 boys) was determined as the experimental group, and class 4/D (15 girls, 12 boys) was determined as the control group.

Findings and Interpretations Regarding Quantitative Research Questions

Findings regarding the post-test scores of the experimental and control groups on the financial literacy academic achievement test

The post-test scores of the experimental and control groups on the academic achievement test were compared using an independent samples t-test. Table 5 presents the results of the independent samples t-test for the post-test scores of the experimental and control groups on the academic achievement test.

Table 5. Results of the Independent Samples t-Test Analysis of the Post-Test Scores of the Experimental and Control Groups on the Financial Literacy Academic Achievement Test

	n	$\overline{X}$	ss	sd	t	p	*d*
Experimental Group	27	28.33	5.40	52	6.19	.001*	1.68
Control Group	27	18.11	6.65				

* p > .05

Looking at Table 5, according to the results of the independent samples t-test conducted to determine whether there was a significant difference between the post-test scores of the experimental and control groups on the academic achievement test, a significant difference in favor of the experimental group was found between the experimental group's post-test score mean ($\overline{X}$ = 28.33) and the control group's post-test score mean ($\overline{X}$ = 18.11) (p > .05). The effect size value (Cohen's d) between the experimental and control group averages was found to be 1.68. This value indicates that the difference between the experimental and control group post-test scores has a large effect. According to Pallant (2017), if the effect size value is between 0.2 < d < 0.5, it is considered a "small" effect; if it is between 0.5 < d < 0.8, it is considered a "medium" effect; and if d > 0.8, it is considered a "large" effect.

Findings regarding the financial literacy academic achievement test retention scores of the experimental and control groups

The analysis of the experimental and control groups' financial literacy academic achievement test retention scores is presented in Table 6.

Table 6. Results of the Independent Samples t-Test Analysis of the Experimental and Control Groups' Financial Literacy Academic Achievement Test Retention Test Scores

	n	$\overline{X}$	**ss**	**sd**	**t**	**p**	***d***
Experimental Group	27	27.81	5.77	52	6.53	.001*	1.77
Control Group	27	17.00	6.36				

* p > .05

Table 6 shows the results of the independent samples t-test conducted to determine whether there was a significant difference between the experimental and control groups' Financial Literacy Academic Achievement Test retention test scores. The mean retention test scores for the experimental group were $\overline{X}$ = 27.81, while the mean retention test scores for the control group were$\overline{X}$ = 17.00. A significant difference in favor of the experimental group was found (p > .05). This significant difference indicates that educational digital game-based instruction related to financial literacy has a lasting effect on students' academic achievement in financial literacy. The effect size between the experimental and control group averages was found to be d=1.77. This value indicates that the effect between the experimental and control group retention test scores is large.

Findings regarding the comparison of the post-test and retention test scores for financial literacy academic achievement in the experimental group

A dependent samples t-test was conducted to determine whether there was a significant difference between the scores obtained by the experimental group on the post-test and the scores obtained on the retention test. Table 7 shows the dependent samples t-test results for the experimental group.

Table 7. Results of the Dependent Sample t-Test Analysis for the Experimental Group's Financial Literacy Academic Achievement Test Final Test and Retention Test Scores

		n	$\overline{X}$	**ss**	**sd**	**t**	**p**
Experimental	Retention test	27	27. 81	5.77	26	1.92	.065
Group	Final test	27	28.33	5.40			

p > .05

When Table 7 is examined, no significant difference was found between the mean final test scores of the experimental group students on the financial literacy academic achievement test and the mean scores on the retention test (p > .05). In the achievement test conducted after the experimental implementation, the average score of the students on the test was 28.33 ($\overline{X}^{-}$ = 28.33), while the average score on the test conducted three weeks after the implementation ended decreased to 27.81 ($\overline{X}$ = 27.81).

Findings regarding the comparison of the control group's financial literacy academic achievement test post-test and retention test scores

A dependent samples t-test was conducted to determine whether there was a significant difference between the control group's scores on the post-test and the retention test. Table 8 shows the results of the dependent samples t-test for the control group.

Table 8. Results of the Paired-Sample t-Test Analysis of the Control Group's Financial Literacy Academic Achievement Test Post-Test and Retention Test Scores

		n	$\overline{X}$	ss	sd	t	p	*d*
Control Group	Retention test	27	17.00	6.36	26	4.50	.001*	.17
	Final test	27	18.11	6.65				

* p > .05

When Table 8 is examined, a significant difference was found between the mean final test scores of the financial literacy academic achievement test administered to the control group students and the mean retention test scores (p > .05). In the academic achievement test conducted after the experimental implementation, the average score of the students on the test was 18.11 ($\overline{X}$), while the average score on the test conducted three weeks after the implementation decreased to 17.00 ($\overline{X}$). The effect size value between the mean scores of the post-test and the retention test was found to be d=.17. This value indicates that the retention level has a statistically low effect.

Findings regarding the comparison of the post-test scores obtained by the experimental and control groups on the overall scale and sub-dimensions of the Financial Literacy Scale for Elementary School Students

The final test scores obtained by the experimental and control groups on the overall scale and sub-dimensions of the Financial Literacy Scale for Elementary School Students were compared using an independent groups t-test. The results of this analysis are presented in Table 9.

Table 9. Results of the Independent Samples t-Test Analysis of the Final Test Scores on the Overall Scale and Subdimensions of the Financial Literacy Scale for Elementary School Students in the Experimental and Control Groups

Scale	Groups	n	$\overline{X}$	ss	sd	t	p	*d*
Overall Scale	Experimental Group	27	47.44	2.48	52	5.36	.001*	1.45
	Control Group	27	43.04	3.46				
Planned Individual	Experimental Group	27	13.37	1.21	52	3.03	.004*	.82
sub-dimension	Control Group	27	12.04	1.93				
Thrifty Individual	Experimental Group	27	11.74	.44	52	4.43	.001*	1.20
Subscale	Control Group	27	10.67	1.17				
Wasteful Individual	Experimental Group	27	22.33	1.54	52	4.41	.001*	1.19
sub-dimension	Control Group	27	20.30	1.83				

* p > .05

Looking at Table 9, according to the results of the independent sample t-test conducted to check whether there was a significant difference between the post-test scores of the experimental and control groups on the overall scale and sub-dimensions of the Financial Literacy Scale for Primary School Students (Özer & Ersoy, 2022), the mean post-test scores of the experimental group on the overall scale were $\overline{X}$ = 47.44, while the control group's mean post-test score was$\overline{X}$ = 43.04. This indicates a significant difference in favor of the experimental group (p > .05). Furthermore, when the experimental group's subscale post-test scores were compared with the control group's subscale scores, a significant difference in favor of the experimental group was found in all subscales. This significant difference indicates that educational digital game-based instruction related to financial literacy significantly increased students' attitudes and behaviors towards financial literacy. The effect size value between the experimental and control group averages was found to be d = 1.45 for the scale as a whole. The effect

size values for the scale subdimensions were found to be *d*=0.82 for the *Planned Individual* subdimension, *d*=1.20 for the *Thrifty Individual* subdimension, and *d*=1.19 for the *Wasteful Individual* subdimension. The results for the overall scale and subdimensions show that the difference between the final test scores of the experimental and control groups has a large effect.

Information regarding the comparison of pre-test, post-test, and retention test scores obtained from the scale overall and sub-dimensions of the Financial Literacy Scale for Elementary School Students in the experimental and control groups

Under this heading, graphical data comparing the pre-test, post-test, and retention test scores of the scale's overall score and sub-dimensions of the Financial Literacy Scale for Elementary School Students (Özer & Ersoy, 2022) applied to the experimental and control groups are presented in Figure 6.

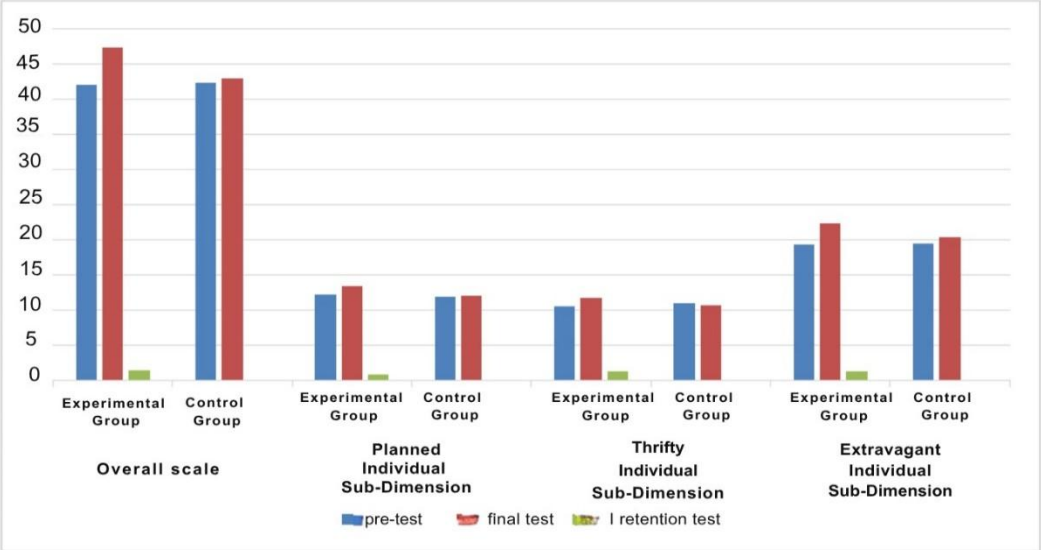


Figure 6. Graphical Representation of the Comparison of Pre-test, Post-test, and Retention Test Scores Obtained by the Experimental and Control Groups on the Overall Scale and Sub-dimensions of the Financial Literacy Scale for Elementary School Students

When examining Figure 6, the analysis of the pre-test and post-test scores of the experimental and control groups shows that the experimental group demonstrated a significant increase across the entire scale and each sub-dimension. In particular, the improvements achieved in the *Planned Individual*, *Thrifty Individual*, and *Wasteful Individual* sub-dimensions reveal the impact of the implementation on the students. The effect size values, evaluated as a persistence test, also show that these gains are not only short-term but also sustainable.

Findings and Interpretations Related to the Qualitative Research Question

The findings obtained from the qualitative data in the study were analyzed under two headings: findings related to students' opinions on the use of educational digital game-based teaching activities in imparting financial literacy skills and findings related to the imparting of financial literacy skills. The analysis of the data obtained regarding the participants' thoughts on teaching with educational digital games revealed the theme "Educational Digital Game-Based Learning Experience" and six sub-themes related to this theme. The analysis of the data regarding the acquisition of financial literacy skills revealed the theme "Financial Literacy Awareness" and nine sub-themes under this main theme. Figure 7 provides information about these themes and sub-themes.

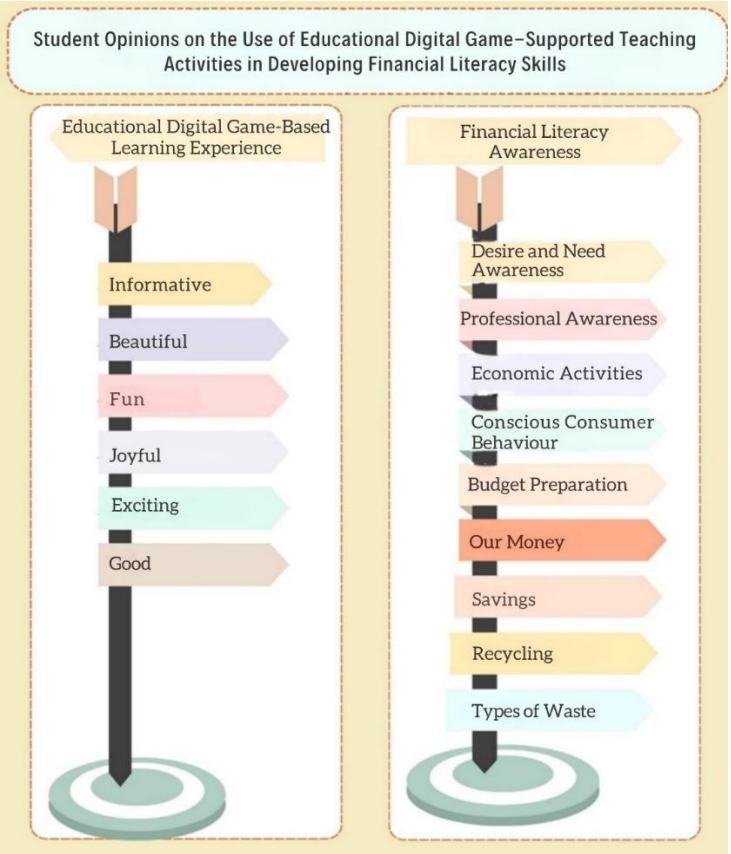


Figure 7. Themes and Sub-themes Identified Through Qualitative Data Analysis

Figure 7 presents the themes and sub-themes that emerged based on the findings obtained from the analysis of the qualitative data of the study. The sub-theme "informative" obtained from the findings of the qualitative data coincides with the findings of increased academic success obtained from the quantitative data of the study. Within this theme, findings regarding students' acquisition of knowledge on the subject as a result of courses conducted with educational digital games are presented. Students stated that they understood the subjects better and were more eager to learn in digital game-supported courses. Regarding this situation, the student named Ela stated: "*We understand better when we play games in our courses. We concentrate more on our courses.*" Within the theme of conscious consumer behavior, the findings obtained regarding students' development of skills in recognizing their rights in consumption processes, conscious shopping, and critical evaluation as a result of courses conducted with educational digital games are presented. For example, a student named Yaren said: "*Today in class, I learned that we should prepare a shopping list before going shopping and ask for a receipt or invoice when leaving. We should buy products with the TSE* [Turkish Standards Institute] *stamp and that are not past their expiration date. I learned that the consumer helpline number is 175*." Within the scope of the savings theme, the findings obtained regarding the students' awareness of efficient use of resources, avoidance of waste, and energy savings in daily life as a result of course conducted with educational digital games are presented. The student named Gökçe said: "*We learned about saving, we learned about saving money. If we have extra money, we will put it in our piggy bank. We learned about the types of waste. We learned about recycling. We don't need to buy a new notebook when we have one. We should recycle our used notebooks.*" Within the scope of the recycling theme, courses conducted with educational digital games included findings that students gained awareness about the importance of recycling, learned to distinguish between recyclable and non-recyclable materials, and developed environmental responsibility. The student named Tarık stated: "*In class, we learned that bread waste is food waste. We learned which foods can be recycled. We learned that meatballs can be made from stale bread. We learned about saving and waste. We learned that glass, metal, cardboard, and paper can be recycled.*" Furthermore, the sub-themes of conscious consumer behavior, saving, and recycling obtained from the qualitative data are consistent with the quantitative findings of the study, which are the awareness of planned and thrifty behavior. These findings reveal that the qualitative and quantitative data overlap.

Findings related to educational digital game-based learning experiences

Students participating in the study were asked what they thought about learning courses through educational digital games. Regarding the teaching of social studies courses through educational digital games, a student named Yaren stated her thoughts as follows: "*It was very nice; each game had a different concept. I learned more than I did when we covered the topic in the usual way. Every game was very nice. I had a lot of fun*." Similarly, a student named Ela expressed her thoughts as follows: *"Social studies classes are very enjoyable when we play games. We understand better when we play games during our courses. Our concentration in class increases*." Looking at the students' statements, they indicated that they understood the course better and learned a lot in courses supported by educational digital games, that they had a lot of fun in class, and that the courses were more enjoyable. One of the most important features of educational digital games is that they make learning fun. A student named Ömer expressed how much he enjoyed learning in class: *"I had so much fun today and I was very happy. I had a lot of fun in the production, consumption, and distribution game. I thought I lost in the raw material and product game, but I won and I was very happy. Even though I lost the next game, I was still very happy*." A student named Nida also expressed her happiness, saying, "*The games we played were very nice, I was very happy*." A student named Mustafa expressed his love for the class, saying, "*I loved it very much, I'm sure my other friends loved it too. I wish we could do it every day because it's a great way to pass the time*." A student named Nida said, "*I was very excited in class, all the games were beautiful. I'm sure my friends loved it too. I had a great time today and was sad when it ended*." Mehmet explained his appreciation for the games: "*I really liked the airplane game, and everyone really liked the quiz game; I think it was very nice*."

Findings related to fostering financial literacy awareness

Regarding the acquisition of financial literacy skills through educational digital games, a student named Yaren expressed her thoughts, saying, *"I learned to distinguish between* wishes *and needs. I learned that we must first meet our needs. What are our needs? What are our wishes? I learned these things and had a lot of fun."* From the student's statement, it is understood that she learned what wishes and needs are and grasped which one to prioritize. Similarly, a student named Tarık stated, "*In social studies class, we learned about wishes, needs, and social needs. I learned* that *going to the movies, theatre, and reading books are social needs."* He also gave examples to explain what a social need is. The student named Gökçe said the following about career awareness and economic activities: *"We learned about raw materials and economic activities. We learned about production, distribution, and consumption, and we learned about different professions. We should choose professions that suit us. Buying bread from the bakery is consumption. We should choose professions that suit us, not those with high salaries."* The student emphasized that individuals should choose professions that suit them. Similarly, a student named Yaren said: "*We learned about products and raw materials. We learned about production, consumption, and distribution. We learned about economic activities: agriculture, animal husbandry, and industry. We learned about professions: products and services. We learned about professions. We learned why we should choose a particular profession. As an example of a product, we can take a cup or a table; as an example of raw materials, we can take iron or steel; as an example of production, we can take the production of wheat; as an example of distribution, we can take the distribution of bread; and as an example of consumption, we can take eating toast."* In this way, she explained the importance of choosing a profession and what economic activities are. The student named Esra talked about conscious consumer behavior: *"Today, we learned about recycling. We learned to pay attention to the expiration date* (*ED*) and *production date of foods, and we learned about recyclable materials. We learned that we should buy local products. We learned about products with TSE quality certificates, we learned to prepare a shopping list, we learned about the consumer arbitration board, and we learned that we should save our extra pocket money."*

The student named Nida said, "*We learned about income, expenses, budget, and coins. I learned how to balance my budget. We learned about Turkish coins, and I learned that receiving a salary is income. I learned that buying a coat for myself in winter is an expense, and I learned that fifty kuruş is a coin*." The student named Yaren said, *"Today in class, we learned about the types of waste, that we can borrow when our income is less than our expenses, that we can save when our expenses are less than our income, and about recyclable and non-recyclable materials. We learned that water waste, bread waste, clothing waste, paper waste, and time waste are all types of waste, and so much more. Glass, plastic, paper...* etc*."*

Discussion, Conclusion, and Recommendations

Discussion and Conclusion

In this study, which examined the effect of educational digital games on the acquisition of financial literacy skills in social studies courses, since there was no significant pre-test difference between the groups in the primary school where the implementation would be carried out, the experimental and control groups were determined by drawing lots from the 4th grade branches. After the 6-week implementation process, an academic achievement test was administered to both groups as a post-test. As a result of the 6-week experimental implementation, a statistically significant increase was observed in the mean success and attitude scores of the experimental group students.

One of the important findings of this study is that there was a statistically significant increase in the financial literacy achievement and attitude scores of students in the 4th grade social studies course taught with educational digital games in favor of the experimental group. In other words, it has been revealed that the financial literacy knowledge and attitudes of elementary school students can be developed in social studies courses taught with educational digital games. In the literature, there are studies on the effect of educational digital games on academic achievement in different subjects at the primary education level. These studies include science (Ağırgöl et al., 2022; Bağ, 2020; Ivgin & Akçay, 2024), mathematics (Boussaha et al., 2025), social studies (Doğan & Koç, 2017; Erkan, 2019; Koka, 2018), and life skills (Kaynar, 2020). This research was conducted in social studies. However, there are also research results in the literature indicating that educational digital games increase academic achievement in different subjects. For example, Ivgin and Akçay (2024) determined that the combined use of educational games and educational digital games significantly increased the science achievement of 5th grade students, but that the use of the games separately had no effect on academic achievement. The finding that the use of educational digital games alone increases academic achievement partially coincides with the results of Ivgin and Akçay's (2024) study. In the research conducted by Boussaha et al. (2025) with 2nd and 3rd grade elementary school students in mathematics courses, the effect of educational digital games on the development of students' arithmetic gains was experimentally demonstrated. Boussaha et al., (2025) conducted their study with 2nd and 3rd grade elementary school students in mathematics course, while this study was conducted with 4th grade students in social studies course. As a result, using educational digital games, Boussaha et al., (2025) determined an increase in arithmetic skills in elementary school, and this study determined an increase in financial literacy knowledge and attitudes. According to the results of both studies, it is a common finding that educational digital games statistically demonstrate a significant increase in students' academic achievement. Chuang and Chen (2009) found that computer-based video games showed a significantly high level of performance in elementary school students' knowledge recall and problem-solving skills. This result is consistent with the finding that educational digital games in this study increased students' academic achievement. Bağ (2020) revealed in their research that educational digital games improved 4th grade students' conceptual understanding levels, scientific thinking habits, and argumentation skills in science courses. This result is similar to the result of the conducted research. Ağırgöl et al. (2022) determined in their research that educational digital games in science courses were statistically significantly higher in terms of both academic achievement and retention of information. There are various similarities and differences between this study and the research conducted. Both studies show similarities in terms of increased academic achievement and lasting learning. Doğan and Koç (2017) found that educational digital games in social studies courses were effective in improving academic achievement. In this respect, the results of both studies coincide. Erkan (2019) found that educational digital games in social studies classes showed a significant increase in academic achievement, but this result was not statistically significant between groups. The result of this study is similar to the result of the research conducted in terms of the increase in academic achievement, but differs in terms of statistical significance. Koka (2018) found that computer-based educational games in 4th grade social studies classes increased students' academic achievement levels and learning retention. The results of both studies coincide in terms of increasing academic achievement and ensuring learning retention.

One of the important findings of the study is that the digital tools used in the development of educational digital games are consistent with similar studies conducted in the field. In this study, Web 2.0 tools such as Wordwall and Scratch were used in the development of digital games; in addition, ready-made recycling-themed games available on the Cevkococuk.org website were also used. Boussaha et al., (2025) used a digital platform called GAME-calcul, which they developed themselves, in their research. The study conducted by Ivgin and Akçay (2024) included both educational and digital games; various ready-made games (Catch a Mole, Space Attack, Space Sale, Recyclebuse, Falling) were used as digital games, along with content from the Cevkococuk.org website. Chuang and Chen (2009) included the three-dimensional digital game Fire Captain in their study on ready-made commercial games. Bağ (2020) used a game he developed himself called Unity in his study. PowerPoint was used in the preparation of educational digital games in the research by Ağırgöl et al. (2022). Digital games were developed in the Adobe Captivate program by Doğan and Koç (2017). Scratch-based digital games were included in the study by Erkan (2019). PowerPoint was used in the preparation of educational digital games in the research by Koka (2018). As a result, the literature review revealed that researchers either developed the games themselves, preferred ready-made commercial games, used both together, or had experts develop the games for them. Most of the digital games used in this study were originally developed by the researcher; in addition, three games were implemented through the Cevkococuk.org website.

In addition to their effects on academic achievement, educational digital games have also been observed to have a significant impact on students' attitudes. In this context, various studies examining the effect of educational digital games on students' attitudes are found in the literature. The research conducted by Dursun (2024) examined the effect of educational digital games on the attitudes of second-grade elementary school students towards learning mathematics. The research findings reveal that the use of educational digital games in mathematics courses has a significant effect on students' attitudes compared to traditional teaching methods. In the experimental group, a significant increase in students' attitudes towards mathematics courses was observed with the use of educational digital games, while in the control group, a decrease in attitude was observed. In a study conducted by Çankaya and Karamete (2008), students' attitudes towards educational digital games in mathematics courses were examined. The results of the study revealed that students had a positive attitude towards educational computer games. A study conducted by Erkan (2019) examined the effect of educational digital games on students' attitudes in social studies. While there was no significant difference between the pre-test and post-test scores of the experimental and control groups, a significant difference in favor of the experimental group emerged between the post-test scores of the experimental and control groups. In their research, Ağırgöl et al. (2022) examined the effect of educational digital games on students' attitudes in science courses and found no significant difference between the experimental and control groups. These findings are similar to the results of our study and show that the effect of educational digital games on students' attitudes may vary depending on factors such as the duration of implementation and content.

One of the qualitative findings of the study is that students reported understanding financial literacy topics better in social studies courses taught using educational digital games. Students also stated that they found the courses very enjoyable and liked them very much. Similar qualitative findings to those of this study have been reported in primary education research in the field. For example, in his study with 6th grade students, Aksoy (2014) reported that students involved in digital game-based mathematics instruction indicated that their interest in learning had increased, that the courses were enjoyable, and that learning had become easier. Avcı et al., (2009) stated that students found courses taught with educational computer games more enjoyable, their participation in the courses increased, and their learning motivation was positively affected. In a study evaluating educational digital games by Bozkurt (2013), it was emphasized that educational digital games attract students' attention, make the learning process effective and enjoyable, and increase students' interest and participation in the course. The qualitative results of the study show that financial literacy topics taught through educational digital games in social studies courses are better understood, more enjoyable, and that

students like the courses, which are informative. These results are consistent with the findings in the literature (Aksoy, 2014; Avcı et al., 2009) and evaluations (Bozkurt, 2013).

In the study, a retention test was administered to the experimental and control groups three weeks after the end of the implementation to understand the effect of educational digital games on the retention of academic achievement, in addition to their effect on academic achievement. The analysis revealed that there was no significant difference between the post-test and retention test scores for the experimental group, while there was a significant difference between the post-test and retention test scores for the control group. This result can be interpreted as indicating that teaching with educational digital games has a positive effect on the retention of learning. This result is consistent with the findings of similar studies examining the effect of teaching with educational digital games on the retention of academic achievement. For example, educational games in English teaching (Donmuş, 2012), teaching English vocabulary (Günel, 2019), teaching the learning domain of Systems in Our Body and Health in science courses (Ağırgöl, 2020), and other disciplines related to educational digital games (Bayat et al., 2014; Karamustafaoğlu & Kaya, 2013; Yapıcı & Karakoyun, 2017; Yiğit, 2007) have found that they have an effect on the permanence of academic success. In this respect, it is understood that the permanence test results of the study are consistent with the results of some studies in the field.

Another important finding of the study is that educational digital games significantly improved students' financial literacy compared to the control group. For this purpose, the Financial Literacy Scale for Elementary School Students (Özer & Ersoy, 2022) was administered as a pre-test to the experimental and control group students, and no significant difference was found between the experimental and control groups. After the 6-week implementation period, the scale was administered as a post-test, and the analysis revealed a significant difference between the experimental group's pre-test and post-test scores on the scale overall and its sub-dimensions. In the control group, a significant difference was found only in the "sub-dimension". Based on this result, it appears that teaching with educational digital games is more effective in developing financial literacy. Among the qualitative results of the study, the students expressed their views on the financial literacy topics covered in educational digital games, including: they were able to better understand their desires and needs, learned what to consider when choosing a profession, learned about economic activities, conscious consumer behavior, how to save, what a budget is and how to make one, and learned about recycling and types of waste. When the quantitative and qualitative results achieved in this context regarding financial literacy are examined together, it is understood that the results are similar and overlapping, and that teaching through educational digital games is confirmed to be effective in imparting financial literacy skills. There are studies on financial literacy in social studies and other subjects in the literature. In Arıkan's (2021) study, which examined the effect of the context-based learning approach in social studies on financial literacy, it was concluded that students learned the importance of budgeting, the distinction between wishes and needs, and how to spend and save. Similarly, Çelikten (2020), in his study examining the impact of the Marmara Financial Literacy Program he developed on financial literacy, determined that students learned what their wishes and needs were; the importance of budgeting, waste, saving, planning, and conscious consumer behavior. The results of the studies by Arıkan (2021) and Çelikten (2020) are particularly consistent with the qualitative results of the present study.

Blue et al. (2018) found that financial education provided in mathematics classes resulted in positive developments in the financial decision-making processes of fourth-grade students. This result is similar to the findings of the present study, particularly in terms of identifying students' wishes and needs and determining that needs should be met first. Kandak and Mertol (2021) found that before receiving financial education, most elementary and middle school students did not know what taxes were or what receipts and invoices were used for. After the education provided, it was concluded that students' awareness of taxes increased and that they learned concepts such as receipts and invoices. This result is similar to the statements in the conducted research that students learned concepts such as receipts and invoices.

There are studies in the literature on financial literacy with elementary school students (Çarıkçı, 2019; Çelikten, 2020; Şimşek, 2023). Since this study attempts to impart financial literacy skills to elementary school students through educational digital games, it is thought that the research results could contribute to the literature. The study is limited to students in two 4th grade classes at an elementary school in a district center, examining the effect of teaching with educational digital games in social studies courses on financial literacy. At the same time, the activities prepared to impart financial literacy are based on the learning outcomes of a learning domain included in the 4th grade Social Studies Course Curriculum. By overcoming these limitations, learning outcomes based on retention and academic achievement can be observed in other grade levels and different learning domains.

Recommendations

This study investigated the use of educational digital games to develop financial literacy skills. Given the limited number of similar studies in this field, it can be said that this research fills an important gap. The results of the study show that educational digital games are an effective tool for improving the financial literacy levels of elementary school students. Considering the limitations of this study, it is important for future research to address this gap by conducting studies at different grade levels and in different learning domains. Furthermore, longitudinal studies can be conducted to track the sustainability of financial literacy skills and their impact on academic achievement over a longer period. Again, more comprehensive studies are needed to investigate the role of different educational methods, primarily educational digital games, in financial literacy education. Such studies can assist educators and policymakers in determining the most effective strategies for financial literacy education. During the research, it was observed that students learned while having fun playing educational digital games. Educational digital games designed for courses can be encouraged to be played by students at home. Teachers can use Web 2.0 tools and free game platforms to develop educational digital games and improve their teaching, especially in imparting skills in different subjects.

References

Adalar, H. (2019). *Sosyal bilgiler öğretmen eğitiminde finansal okuryazarlık ve öğretimi: Bir model önerisi* (Thesis No. 537370) [Doctoral dissertation, Gazi University]. Council of Higher Education National Thesis Center.

Ağırgöl, M. (2020). *Fen bilgisi öğretiminde eğitsel dijital oyun kullanımının öğrenci akademik başarısına bilgi kalıcılığına ve tutumuna etkisi* (Thesis No. 639945) [Master's thesis, Binali Yıldırım University]. Council of Higher Education National Thesis Center.

Ağırgöl, M., Kara, E., & Dönel Akgül, G. (2022). Eğitsel dijital oyunlarla işlenen fen bilgisi dersinin öğrencilerin bilgilerinin kalıcılığına, akademik başarısına ve tutumuna etkisi. *Uluslararası Bilim ve Eğitim Dergisi, 5*(3), 157-176. https://doi.org/10.47477/ubed.1063920

Akengin, H., & Ersoy, F. (2015). Sosyal bilgiler eğitiminde mekansal öğrenme ortamlarının tarihçesi. In R. Sever & E. Koçoğlu (Eds.), *Sosyal bilgiler eğitiminde mekansal öğrenme ortamları* (pp. 17-42). Pegem.

Akgün, E., Nuhoğlu, P., Tüzün, H., Kaya, G., & Çınar, M. (2011). Bir eğitsel oyun tasarımı modelinin geliştirilmesi. *Eğitim Teknolojisi Kuram ve Uygulama*, *1*(1), 41-61.

Akhan, N. E. (2009). Sosyal bilgilerde ekonomi konularının öğretimi. In M. Safran (Ed.), *Sosyal bilgiler öğretimi* (pp. 180-195). Pegem.

Aksoy, N. C. (2014). *Dijital oyun tabanlı matematik öğretiminin ortaokul 6. sınıf öğrencilerinin başarılarına, başarı güdüsü, öz-yeterlik ve tutum özelliklerine etkisi* (Thesis No. 356671) [Doctoral dissertation, Gazi University]. Council of Higher Education National Thesis Center.

Amagir, A., Maassen van den Brink, H., Groot, W., & Wilschut, A. (2022). SaveWise: The impact of a real-life financial education program for ninth grade students in the Netherlands. *Journal of Behavioral and Experimental Finance*, (33), 100605. https://doi.org/10.1016/j.jbef.2021.100605

Arıkan, İ. (2021). *Sosyal bilgiler dersinde bağlam temelli öğrenme yaklaşımının öğrencilerin akademik başarısına, finansal okuryazarlık becerilerine ve öğrenmenin kalıcılığına etkisi* (Thesis No. 683469) [Doctoral dissertation, Fırat University]. Council of Higher Education National Thesis Center.

Avcı, Ü., Sert, G., Özdinç, F., & Tüzün, H. (2009). Eğitsel bilgisayar oyunlarının bilişim teknolojileri dersindeki kullanım etkileri. In P. Akar, B. Akkoyunlu, A. Altun, M. Erdem, S. S. Seferoğlu, Y. Koçak Usluel, H. Tüzün, A. Özkök, & H. Yurdugül (Eds.), *9. International Educational Technology Conference* (pp. 64-69). Hacettepe Üniversitesi.

Bağ, H. (2020). *Eğitsel bir dijital oyun yardımıyla kavramsal anlama düzeyinin, bilimsel düşünme alışkanlıklarının ve argümantasyon becerilerinin gelişiminin incelenmesi* (Thesis No. 626596) [Doctoral dissertation, Trabzon University]. Council of Higher Education National Thesis Center.

Baloğlu Uğurlu, E. (2012). Sosyal bilgiler eğitiminde teknoloji araçlarının kullanımı. In M. Safran (Ed.), *Sosyal bilgiler öğretimi* (pp. 243-265). Pegem.

Bayat, S., Kılıçarslan, H., & Şentürk, Ş. (2014). Fen ve teknoloji dersinde eğitsel oyunların yedinci sınıf öğrencilerinin akademik başarısına etkisinin incelenmesi. *Abant İzzet Baysal Üniversitesi Eğitim Fakültesi Dergisi*, *14*(2), 204-216. https://doi.org/10.17240/aibuefd.2014.14.2-5000091535

Baykul, Y. (2015). *Eğitimde ve psikolojide ölçme: Klasik test teorisi ve uygulaması*. Pegem.

Bayrakdaroğlu, A., & Bilge, S. (2018). Kadınlarda finansal eğitimin finansal okur-yazarlık üzerindeki etkisine yönelik bir araştırma [Special issue]. *Sosyal ve Beşeri Bilimler Araştırmaları Dergisi*, *19*(42), 97-120.

Blue, L. E., O'Brien, M., & Makar, K. (2018). Exploring the classroom practices that may enable a compassionate approach to financial literacy education. *Mathematics Education Research Journal,* (30)**,** 143-164. https://doi.org/10.1007/s13394-017-0223-5

Boussaha, K., Drissi, S., Rahab, A., Touhami, K. K., & Boudraa, C. E. A. T. (2025). Towards a new generation of digital games designed with the basics of psychological theories to improve primary school pupils' psychological and technical skills in learning arithmetic. *Entertainment Computing*, (53), 100927. https://doi.org/10.1016/j.entcom.2025.100927

Bozkurt, Ö. F. (2013). Eğitsel dijital oyun örnekleri. In M. A. Ocak (Ed.), *Eğitsel dijital oyunlar: Kuram, tasarım ve uygulama* (pp. 197-214). Pegem.

Büyüköztürk, Ş. (2014). *Sosyal bilimler için veri analizi el kitabı* (19th ed.). Pegem.

Cheng, M. T., She, H. C., & Annetta, L. A. (2015). Game immersion experience: Its hierarchical structure and impact on game-based science learning. *Journal of Computer Assisted Learning*, *31*(3), 232-253. https://doi.org/10.1111/jcal.12066

Chuang, T.-Y., & Chen, W. F. (2009). Effect of computer-based video games on children: An experimental study. *Educational Technology & Society*, *12*(2), 1-10.

Clark, D. B., Tanner-Smith, E. E., & Killingsworth, S. S. (2016). Digital games, design, and learning: A systematic review and meta-analysis. *Review of Educational Research*, *86*(1), 79-122. https://doi.org/10.3102/0034654315582065

Coleman, T. E., & Money, A. G. (2019). Student-centered digital game–based learning: A conceptual framework and survey of the state of the art. *Higher Education*, (79), 1-43. https://doi.org/10.1007/s10734-019-00417-0

Creswell, J. W., & Plano Clark, V. L. (2018). *Designing and conducting mixed methods research* (3rd ed.). Sage.

Çankaya, S., & Karamete, A. (2008). Eğitsel bilgisayar oyunlarının öğrencilerin matematik dersine ve eğitsel bilgisayar oyunlarına yönelik tutumlarına etkisi. *Mersin Üniversitesi Eğitim Fakültesi Dergisi*, *4*(2), 115-127.

Çarıkçı, O. (2019). İlkokul dördüncü sınıf öğrencilerinin finansal farkındalık düzeyinin belirlenmesine yönelik bir araştırma. *Akdeniz Üniversitesi Sosyal Bilimler Enstitüsü Dergisi*, (5), 133-157.

Çelikten, L. (2020). *Marmara Finansal Okuryazarlık Programı'nın geliştirilmesi ve ilkokul 4. sınıf öğrencilerinin finansal okuryazarlık becerilerine etkisinin incelenmesi* (Thesis No. 655535) [Doctoral dissertation, Marmara University]. Council of Higher Education National Thesis Center.

Dai, C. P., Ke, F., & Pan, Y. (2022). Narrative-supported math problem solving in digital game-based learning. *Educational Technology Research and Development*, *70*(4), 1261-1281. https://doi.org/10.1007/s11423-022-10129-5

Doğan, E., & Koç, H. (2017). Sosyal bilgiler dersinde deprem konusunun dijital oyunla öğretiminin akademik başarıya etkisi. *Uluslararası Türk Eğitim Bilimleri Dergisi*, *5*(8), 90-100. https://dergipark.org.tr/tr/download/article-file/408585

Donmuş, V. (2012). *İngilizce öğrenmede eğitsel bilgisayar oyunu kullanmanın erişiye kalıcılığa ve motivasyona etkisi* (Thesis No. 323364) [Master's thesis, Fırat University]. Council of Higher Education National Thesis Center.

Dursun, Z. B. (2024). *İlkokulda eğitsel dijital oyunun matematik tutumuna etkisi* (Thesis No. 871403) [Master's thesis, Necmettin Erbakan University]. Council of Higher Education National Thesis Center.

Er, B., & Çetintaş, M. (2018). Çalışanlarda finansal okuryazarlık: finansal bilgi, tutum ve davranışların değerlendirildiği bir araştırma. *Mersin Üniversitesi Sosyal Bilimler Enstitüsü Dergisi*, *1*(2), 61-71.

Erkan, A. (2019). *İlkokul 4. sınıf sosyal bilgiler dersinde kullanılan eğitsel oyun ve dijital oyun öğretiminin öğrencilerin başarı ve tutumlarına etkisi* (Thesis No. 567523) [Master's thesis, Binali Yıldırım University]. Council of Higher Education National Thesis Center.

Erlandson, D. A., Harris, E. L., Skipper, B. L., & Allen, S. D. (1993). *Doing naturalistic inquiry: A guide to methods*. Sage.

Fadda, D., Pellegrini, M., Vivanet, G., & Zandonella Callegher, C. (2022). Effects of digital games on student motivation in mathematics: A meta-analysis in K-12. *Journal of Computer Assisted Learning*, *38*(1), 304-325. https://doi.org/10.1111/jcal.12618

Farris, P. J. (2004). Social studies and integrated instruction: A look at social studies. In P. J. Farris (Ed.), *Elementary and middle school social studies: An interdisciplinary, multicultural approach* (4th ed., pp. 2-43). McGraw-Hill.

Fidancı, H. (2021). *Lise öğrencilerinin finansal okuryazarlık algılarının farklı değişkenlere göre incelenmesi* (Thesis No. 694125) [Master's thesis, Erciyes University]. Council of Higher Education National Thesis Center.

Gömleksiz, M., & Erkan, S. (2010). *Eğitimde ölçme ve değerlendirme* (2nd ed.). Nobel

Guest, G., & Fleming, P. (2015). *Mixed methods research*. In G. Guest, E. E. Namey, G. Guest, & P. Fleming (Eds.), *Public health research methods* (pp. 581-614). Sage. https://doi.org/10.4135/9781483398839.n19

Günel, E. (2019). *Gerçek çevre tabanlı dijital eğitsel oyunların İngilizce kelime bilgisi ile kalıcılığı üzerine etkilerinin incelenmesi* (Thesis No. 574119) [Master's thesis, Bolu Abant İzzet Baysal University]. Council of Higher Education National Thesis Center.

Güvenç, H. (2017). Öğretim programlarımızda finansal okuryazarlık. *İlköğretim Online, 16*(3), 935-948. https://doi.org/10.17051/ilkonline.2017.330233

Ivgin, A. B., & Akcay, H. (2024) The impact of using educational and digital games on middle school students science achievement. *International Journal of Technology in Education, 7*(3), 386-416. https://doi.org/10.46328/ijte.781

Kandak, M., & Mertol, G. (2021). Eğitim düzeyi ile vergi gelirleri ilişkisi: İlkokul ve ortaokul öğrencileriyle ampirik bir çalışma. *Eğitim Bilim ve Araştırma Dergisi, 2*(1), 9-24.

Karamustafaoğlu, O., & Kaya, M. (2013). Eğitsel oyunlarla 'yansıma ve aynalar' konusunun öğretimi: Yansımalı konu örneği. *Araştırma Temelli Etkinlik Dergisi, 3*(2), 41-49.

Kaynar, B. (2020). *Eğitsel ve dijital oyun tabanlı etkinliklerin hayat bilgisi dersindeki akademik başarı, tutum ve kalıcılığa etkisi* (Thesis No. 653030) [Master's thesis, Atatürk University]. Council of Higher Education National Thesis Center.

Koka, V. (2018). *Sosyal bilgiler dersinde kullanılan bilgisayar destekli eğitsel oyunların öğrencilerin ders başarısına olan etkisi* (Thesis No. 505026) [Master's thesis, İnönü University]. Council of Higher Education National Thesis Center.

Lusardi, A. (2019). Financial literacy and the need for financial education: Evidence and implications. *Swiss Journal of Economics and Statistics*, *155*(1), 1. https://doi.org/10.1186/s41937-019-0027-5

Mandouit, L., & Hattie, J. (2023). Revisiting "The Power of Feedback" from the perspective of the learner. *Learning and Instruction*, *84*(6), 101718. https//doi.org/10.1016/j.learninstruc.2022.101718

OECD. (2012). *Supplementary questions: Optional survey questions for the OECD INFE Financial Literacy Core Questionnaire*. https://www.oecd.org/finance/financial-education/49878153.pdf

OECD. (2020). *PISA 2018 assessment and analytical framework.* https://doi.org/10.1787/19963777

Özdamar, K. (2016). *Eğitim, sağlık ve davranış bilimlerinde ölçek ve test geliştirme yapısal eşitlik modellemesi.* Nisan Kitabevi.

Özer, S., & Ersoy, A. (2022). Developing a financial literacy scale for primary school students: A validity and reliability study. *International Journal of Psychology and Educational Studies*, *9*(4), 1197-1207. https://doi.org/10.52380/ijpes.2022.9.4.817

Öztürk, M. E. (2021). *TRB1 bölgesi üniversite öğrencilerinin finansal okuryazarlık düzeylerinin karşılaştırılmasına yönelik bir araştırma* (Thesis No. 671885) [Master's thesis, Munzur University]. Council of Higher Education National Thesis Center.

Pallant, J. (2017). *SPSS kullanma kılavuzu: SPSS ile adım adım veri analizi* (S. Balcı & B. Ahi, Trans.). Anı.

Patton, M. Q. (2014). *Nitel araştırma ve değerlendirme yöntemleri* (M. Bütün & S. B. Demir, Trans.). Pegem.

Sabırlı, Z. E. (2018). *Dijital eğitsel oyunların eğitimde kullanımının farklı değişkenler açısından incelenmesi* (Thesis No. 527313) [Master's thesis, Selçuk University]. Council of Higher Education National Thesis Center.

Samur, Y. (2016). *Dijital oyun tasarımı.* Pusula.

Şimşek, A. (2023). *Sosyal bilgiler dersinde finansal okuryazarlığın geliştirilmesinde drama yönteminin kullanılması: Bir eylem araştırması* (Thesis No. 804733) [Doctoral dissertation, Necmettin Erbakan University]. Council of Higher Education National Thesis Center.

Yapıcı, İ. Ü., & Karakoyun, F. (2017). Biyoloji öğretiminde oyunlaştırma: Kahoot uygulaması örneği. *Turkish Online Journal of Qualitative Inquiry, 8*(4), 396-414. https://doi.org/10.17569/tojqi.335956.

Yıldırım, A., & Şimşek, H. (2013). *Sosyal bilimlerde nitel araştırma yöntemleri* (9th ed.). Seçkin.

Yiğit, A. (2007). *İlköğretim 2.sınıf seviyesinde bilgisayar destekli eğitici matematik oyunlarının başarıya ve kalıcılığa etkisi* (Thesis No. 206271) [Master's thesis, Çukurova University]. Council of Higher Education National Thesis Center.

Zhu, A. Y. F., Yu, C. W. M., & Chou, K. L. (2021). Improving financial literacy in secondary school students: An randomized experiment. *Youth & Society*, *53*(4), 539-562. https://doi.org/10.1177/0044118X19851311

Education and Science

Original Article

Vol 51 (2026) No 226 327-358

The reflection of text structure modifications on reading comprehension: a comparison of authentic and modified texts among 7th grade students

Enes Çinpolat [1], Keziban Tekşan [2], Yunus Emre Uyar [3]

Abstract

This study aims to compare the reading comprehension levels of 7th-grade students in authentic and modified texts. For this purpose, two texts were identified, one being narrative and the other informative. The selected texts were processed according to determined modification criteria, and an effort was made to align them with the students' grade levels. Following this, reading comprehension questions regarding the texts were prepared. Two study groups were utilized during these stages. The first group (n = 22) was included in the text modification process. While the second group (n = 34) was used during the development phase of the reading comprehension questions. After preparing the texts and reading comprehension questions, two groups equivalent in terms of reading comprehension were selected. One group was presented with the modified text, and the other group received the authentic text. Subsequently, reading comprehension questions related to the read texts were directed to these two groups. Thus, the third group in the research (n = 92) participated in the comparison of reading comprehension scores obtained from the authentic and modified texts. The obtained data were analyzed with ANCOVA in the Jamovi software. According to the findings, the contribution of texts modification to reading comprehension is statistically significant. Effect size analyses showed that this contribution was at a medium level (η^2_p = .087) for the narrative text and at a large level (η^2_p = .141) for the informative text.

Keywords

Narrative text
Informative text
Text modification
Reading comprehension
Middle school student

Article Info

Received: 07.03.2024
Accepted: 12.23.2025
Published Online: 04.30.2026

DOI: 10.15390/ES.2026.2508

Introduction

Language is the most important instrument underpinning the educational process, beyond serving merely as a means of communication. Given that teaching and learning activities are largely conducted through language-based resources, it is evident that there is a direct relationship between a student's reading comprehension skills and their general academic achievement (Bayraktar & Durukan, 2016). This situation is particularly pronounced in Turkish courses. Indeed, in these courses, texts serve as the primary material not only for developing basic language skills but also for instilling reading

[1] Ordu University, Faculty of Education, Department of Turkish and Social Sciences Education, Ordu, Türkiye, enescinpolat@gmail.com
[2] Ordu University, Faculty of Education, Department of Turkish and Social Sciences Education, Ordu, Türkiye, kezibanteksan@gmail.com
[3] Ministry of National Education, Samsun, Türkiye, yunus.emre@mit.tc

appreciation, core values, and 21st-century skills (Arslan & Engin, 2019; Derse & Coşkun, 2021; Kurudayıoğlu & Soysal, 2019; Şeref et al., 2020). As stated by the Ministry of National Education (MoNE, 2019), these texts, designed with a thematic approach, constitute the foundation of course activities. Considering the critical role of texts in achieving the objectives of the Turkish course and their importance in the instructional process, the question of which texts students encounter gains significant importance. Therefore, text selection can be considered one of the most crucial and sensitive stages of language teaching. As emphasized in the literature, the prerequisite for deriving maximum benefit from the educational process is that the presented texts are fully aligned with the cognitive and developmental levels of the target audience (Bayraktar & Durukan, 2020; Karaağaç & Alikılıç, 2023).

According to Güneş (2013), in addition to literary texts, original, constructed, and special-purpose texts are utilized in language teaching. Among these, literary texts are defined as products that arouse excitement and admiration in the reader by presenting emotions and thoughts within an aesthetic framework. However, it can be stated that two fundamental problems emerge that complicate the use of these texts in educational settings. The first problem is that many literary works, by virtue of their origin, did not prioritize the reality of the child or pedagogical concerns. Many of these works were written before the development of psychological and pedagogical approaches regarding the concept of childhood. As Şimşek (2007) stated, this situation is the product of an understanding wherein children experience a world specific to adults. Therefore, literary works may not always be suitable for the developmental characteristics of contemporary students in terms of their themes and narrative forms. For instance, even if the protagonist is a child, Ömer Seyfettin's story titled "Kaşağı" contains sections that conflict with the psychological characteristics of a certain age group (such as the direct handling of the phenomenon of death through a child character). Secondly, it is possible that literary texts often possess linguistic and narrative features that may impede understanding for middle school students, thereby potentially fostering negative attitudes toward the act of reading. As noted in Başaran's (2021) study, texts that resonate with the student's life experiences and possess plain and clear language (that is, high readability levels) can positively affect the reading comprehension process. Furthermore, Durukan's (2014) study indicates that texts containing long words and sentences negatively impact students' reading speeds and reading comprehension levels. Although he employed relatively simple language for his era, words exceeding the level of middle school students are encountered in Ömer Seyfettin's stories. Similarly, Muallim Naci's work titled "Ömer'in Çocukluğu" (Ömer's Childhood), despite centering on a child protagonist, contains long and detailed descriptions that freeze the narrative flow, making it significantly difficult for students at this level to follow the text. Disregarding literary texts due to these difficulties might imply forfeiting their indispensable contributions. Foremost among these contributions is that texts reflect the culture in which they were formed in its most natural state. Indeed, as Melanlıoğlu (2008) stated, language teaching is simultaneously culture teaching. In this context, particularly folk literature works, according to Özcan (2008), offer a rich resource harboring a nation's modes of living, thinking, and appreciation spanning centuries. On the other hand, these texts, which have reached artistic competence through long literary processes, are among the most important tools for instilling artistic sensitivity in students, as emphasized by Sever (2018). Briefly stated, despite the national and universal values and aesthetic riches they carry, literary works may surpass students' comprehension levels in terms of language and narrative features (vocabulary, sentence structures, etc.) because they are usually written without observing the principle of suitability for the child. Moreover, certain themes or characters processed in these texts may not be suitable for students' developmental characteristics, and this situation may conflict with the pedagogical and legal frameworks determined by the MoNE (2019). At this point, Duman (2010) states that instead of leaving the student trapped between the unintelligible and the misunderstood, intervening in the text when and to the extent necessary is a viable approach. This situation has brought practices known as text modification (simplification, adaptation, revision, etc.) to the forefront of the literature. The literature treats text modification as a practice that offers significant benefits while also harboring serious limitations. The primary benefit of modification is that it increases reading comprehension levels by reducing the linguistic and structural complexity of texts. This effect is more pronounced on target audiences such as non-native speakers, students with reading difficulties, and children (Gala et al., 2018; Glavas &

Štajner, 2015; Javourey-Drevet et al., 2022; Kim & Snow, 2009). It is also stated that modified (especially simplified) texts make historical or complex subjects more accessible (Durmuş, 2013a) and facilitate students' reading fluency and interaction with the text (Bayraktar & Durukan, 2020; Javourey-Drevet et al., 2022). However, the risks and limitations associated with text modification are also extensively covered in the literature. The fundamental criticism is the potential for the simplification process to damage the text's literary texture, aesthetic value, and the author's original style (Uygur, 2024; Yücelşen, 2014). Changes made with a mechanical approach can transform the work into a soulless text by destroying its rhythm, atmosphere, and the associative richness of its words (Carroli, 2008; Gilmore, 2007). A second limitation is that oversimplification deprives students of the opportunity to cope with complex language structures and develop higher-level thinking skills. A study conducted by Hayes et al. (1996) drew attention to this danger by suggesting that simplification in school textbooks might be related to a decline in students' verbal achievements. Additionally, Dahl et al. (2021) state that students develop more metacognitive and abstract evaluations in complex science texts. This dilemma underscores the necessity of conducting text modification not as an arbitrary and intuitive process, but through an approach with clearly determined principles that preserve both the structural and aesthetic integrity of the text.

It is frequently noted in the literature that the intuitive approach is prevalent in the text modification process (Durmuş, 2013a). Although this approach offers advantages such as flexibility and situational adaptability, it distances the process from objectivity (Uygur, 2024). İsen and Işınsu (2011) explain this situation as publishing houses and editors turning to superficial solutions, such as merely removing unknown words, instead of developing a standard method. Indeed, it is frequently emphasized in the literature that such uncontrolled interventions distort the text's messages, cause it to lose its artistic value, lead to semantic ambiguity and structural inconsistency, and furthermore, create unnatural language input (Bakan, 2012; Durmuş, 2013a; Long & Ross, 1993). Besides this approach, the structural approach, in which the linguistic structures the student will encounter at each stage are predetermined (Aytan et al., 2021), differs from the intuitive approach in this aspect. While the structural approach offers a systematic process based on language structures, it does not always prioritize cohesion and coherence (Uygur, 2024). In contrast, the linguistic approach offers a more in-depth framework for modification by focusing on the semantic integrity and contextual harmony of the text (Bakan-Aktaş & Ay, 2021). Along with these, specific standards, strategies, and techniques for text modification are also suggested in the literature. Guidance resources, such as the MoNE curricula (MoNE, 2006, 2019), address principles of intervention in texts while listing the characteristics that texts to be included in textbooks must possess. These approaches to be applied primarily address interventions made to texts at the lexical and syntactic levels (De Belder & Moens, 2010; Surya et al., 2018). Bölükbaş (2015) classifies the four basic strategies of text modification as preservation (leaving parts suitable for the level as is), modification (simplifying parts not understood), summarizing (shortening while preserving semantic integrity), and deletion (removing details that do not contribute to the semantic flow). Similarly, Yaşar (2019) classifies this process as lexical, syntactic, lexical-syntactic, and formal modification. Yücelşen (2014), on the other hand, includes more micro-level operations such as changing the text title, personal names, and sentence positions in this process under the headings of shortening and editing. As can be seen, the literature defines text modification as a multi-layered and systematic operation extending from the word level to the whole text. In addition to these, Durmuş (2013b) addresses text modification under two basic types: simplification and elaboration. Simplification involves replacing lexical and syntactic structures thought to make understanding difficult for the target audience with simpler, more understandable equivalents. Elaboration, on the other hand, aims to contribute to the target audience's understanding by preserving original words that are difficult to understand in the text and adding synonyms or explanatory expressions to them. Durmuş (2013b) also states that in the text modification process, besides operations that are modification types in themselves such as summarizing, simplification, and elaboration, sub-operations encountered in more than one modification type (such as shortening, splitting, paraphrasing, combining, substitution, dropping, and reduction) also play a role. It is stated that as a result of these operations, reconstructed texts emerge that differ semantically from the source text but are more suitable for the

target audience. Beyond these strategies and techniques, Duman (2010) comprehensively sets forth the philosophy and ethical framework of text modification. According to Duman, preserving the original spirit of the text lies at the center of this process. This principle requires the intervention to maintain the author's style, the sentence structure dominating the text, and its phonetic characteristics. However, the most fundamental and inviolable principle is the preservation of the text's main idea and semantic integrity. As can be seen, the literature offers a body of knowledge regarding the theoretical foundations, applicable techniques, and fundamental principles of text modification.

The presence of text modification in the context of Turkish language teaching has been addressed from various aspects in the literature. A significant portion of studies in this field has focused on descriptively analyzing modification practices in textbooks. Studies such as Bayraktar and Durukan (2016), Tekşan and Çinpolat (2023), and Yücelşen (2014) examined the current situation in the field by revealing, through document analysis, by which methods and to what extent the texts in textbooks were modified. There are also studies that examine the effect of the practice by moving beyond this descriptive framework. Durukan (2014) addressed a topic that can be said to be related to modification by examining the relationship between readability level and reading skills. Bayraktar and Durukan (2020) demonstrated the positive effect of surface structure changes on students' comprehension skills and views employing a mixed-method design, thereby strengthening the foundation of the subject in the context of Turkish education. As can be seen, experimental studies in which authentic and modified texts are directly compared and where the different responses of text types to modification practices are examined empirically are limited. In this context, the current research aims to offer an original and holistic contribution to the literature by comparatively examining the effect of modified forms of different types of texts on students' reading comprehension achievement using quantitative data. In this regard, the current research aims to comparatively examine the contribution of authentic and modified texts on the reading comprehension achievements of 7th-grade students. In the study, students' reading comprehension levels regarding these texts were evaluated by using authentic and modified versions of two different types of texts (narrative and informative). Thus, answers to the following questions were sought within the scope of the research:

1. Is there a contribution of the text modification practice to reading comprehension scores in the narrative text?
2. Is there a contribution of the text modification practice to reading comprehension scores in the informative text?

Method

Design

In this research, a quasi-experimental pre-test post-test control group design was employed. Quasi-experimental designs are research models that permit the examination of the effect of the independent variable, although the controls required by experimental designs cannot be fully established (Creswell, 2020). In the study, two groups consisting of intact classes were determined in their natural states, and random assignment was not performed. Prior to the process of comparing reading comprehension in modified and authentic texts, a reading comprehension test (Ülper et al., 2017) was administered to both groups to evaluate the groups' baseline levels and check their equivalence. Subsequently, the experimental group read the modified text type, while the control group read the authentic text. Following the reading of the texts, reading comprehension questions related to the read texts were administered to both groups, and the scores of the two groups were evaluated. Thus, the differences between the post-test results of the two groups were compared through statistical analyses (see Figure 1).

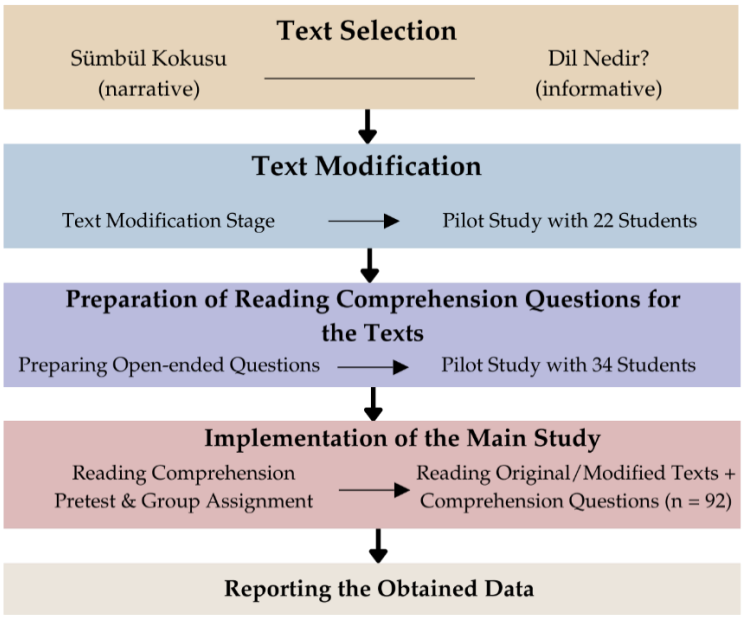


Figure 1. Research Design

Study Group

Three different study groups were used within the scope of the research. The first study group was utilized in the text modification process. The second study group was used during the pilot application and rubric development phase of the open-ended reading comprehension questions. Finally, the third study group was used during the administration phase of the reading comprehension achievement tests for the modified and authentic texts. All study groups consisted of intact classes, and student participation was not mandatory. Data collection in the third group was carried out in stages. Therefore, students who participated in the pre-assessment and were included in one of the stages of the main application remained in the sample. Students who did not participate in the narrative and informative text applications were excluded from the sample. Thus, one student who participated in the narrative text application but not the informative text application, and two students who participated in the informative text application but not the narrative text application, remained in the sample. For this reason, the number of participants is restated in each analysis in the findings section.

The first group consists of 22 students from the 7th grade. The students study at a public school. This study group was used during the modification phase of the texts. This group has no connection with the other groups. The gender distribution of the group is eight boys and fourteen girls. This study group was administered the authentic text and the modified text. They were asked to underline the parts not understood in the authentic and modified texts. Thus, the parts where difficulties in understanding occurred were determined.

The second study group consists of a total of 34 students, also from the 7th grade. These students study at a public school. This group was used to carry out the pilot application of the texts and reading comprehension questions. During the pilot application process, a graded scoring key (rubric) was prepared to ensure the scoring of reading comprehension questions, and example student answers were used in the creation of this key. The reading duration of the authentic text and the modified text was also examined. Thus, final checks were performed on the texts at this stage, and the reading comprehension questions consisting of open-ended questions were finalized.

The third study group, in which modified and authentic texts were compared in terms of reading comprehension, consists of 92 students. 50 girls and 42 boys from four different seventh-grade classes were used in the achievement test. Due to the difficulty of intervening in classes within the scope of the study, matching could not be resorted to in the groups. However, an examination was conducted on whether the reading comprehension achievements and social lives of the classes to be applied showed similar characteristics. First of all, all four classes are located in schools situated in the central

districts of two different metropolitan cities. Through preliminary interviews held with the Turkish teachers of all classes, the opinion emerged that the general achievement statuses of these classes were at a similar level. However, the Reading-Comprehension Test (Ülper et al., 2017), prepared for middle school level students, was used for the purpose of comparing the students' reading comprehension achievements with each other. It was administered to four different groups within one lesson hour and scored by the researchers according to the directive of the measurement tool. The obtained scores were analyzed with the One-Way ANOVA test. According to the obtained result, no significant difference was detected among the reading comprehension achievement scores of the four classes ($F(3,88) = 2.24$, $p = .090$). The mean scores and student numbers of the classes are as follows: 7/C = 19.6 ($n = 22$), 7/D = 21.8 ($n = 25$), 7/E = 18.8 ($n = 18$), 7/K = 20.6 ($n = 27$). Although there was no statistical score difference between the classes, it was observed that the mean scores of two classes exceeded 20 points. For this reason, observing that these two classes should not be in the same group, the modified text and authentic text application groups were selected randomly. Thus, classes 7/D and 7/E were assigned as the modified text group, and classes 7/C and 7/K as the authentic text group. The score difference between the groups formed in this way was also examined with the independent samples t-test. According to the obtained result, no significant difference was observed between the modified ($M = 20.6$, $SD = 4.33$) text group of 43 people and the authentic ($M = 20.2$, $SD = 3.98$) text group of 49 people in terms of reading comprehension scores ($t(90) = 0.483$, $p = .630$; Levene's $F(1, 90) = 0.235$, $p = .629$; Kolmogorov-Smirnov $D(90) = .127$, $p = .101$). In the continuation of the study, these groups were named the modified text group and the authentic text group.

Data Collection Tools

Two texts, one narrative and one informative, and reading comprehension questions were used for data collection in the research. First, work was carried out on the selection of texts. In determining the texts, texts that were suitable for the themes of the middle school Turkish course and had not been used in Turkish textbooks before were researched. For this purpose, primarily the texts used in Turkish textbooks within the last 20 years and the section on the qualities of texts to be included in textbooks of the Turkish Course Curriculum were examined (MoNE, 2019; Tekşan et al., 2023). In line with this examination, texts that frequently appeared in textbooks and were highly likely to have been encountered by students previously were excluded from the scope of the study. Furthermore, children's literature products were also excluded as they would not require modification. Thus, the works titled "Dağa Çıkan Kurt" by Halide Edib Adıvar, "Çağlayanlar" by Ahmet Hikmet Müftüoğlu, "Millî Savaş Hikayeleri" by Yakup Kadri Karaosmanoğlu, "Türkler'in Kültür Kökenleri" by Ergun Candan, "Türk Dil Bilgisi" by Muharrem Ergin, and "Kültür ve Dil" by Mehmet Kaplan were examined by the researchers. These authors and works were selected because they were aligned in a specific direction in terms of thematic and linguistic features. It was aimed that the narrative texts be taken from authors competent in their field in accordance with the National Resistance and Atatürk theme, and the informative texts in accordance with the Our National Culture theme. The fact that these works could not be used directly in textbooks due to their linguistic features, despite being high-quality works, was effective in forming the sample in this study. Following the examination of narrative style, subject, and theme, the text "Sümbül Kokusu" (Scent of Hyacinth), suitable for the National Resistance and Atatürk theme for the narrative text, and the text "Dil Nedir" (What is Language), suitable for the Our National Culture theme for the informative text, were determined. The story titled "Sümbül Kokusu" was written by Ahmet Hikmet Müftüoğlu in 1918. The text "Dil Nedir," selected as the informative type, was chosen from Muharrem Ergin's book titled "Türk Dil Bilgisi." The opinions of three experts in the field of Turkish education were consulted regarding the suitability of these texts for the themes and for modification. Additionally, Durmuş (2013b) stated that among the grounds for text modification, there are cases such as the modification of historical texts and the modification of texts prepared for adults. Based on these opinions and reasons for text modification, the modification and reading comprehension question writing process for the texts whose suitability was approved began. In text modification, the use of only the intuitive method was prevented by utilizing the Turkish course curriculum, sections not understood by students, A Frequency Dictionary of Turkish (Aksan et al., 2016), and A Frequency Dictionary of Contemporary Turkish (Tezcan-Aksu & Adalı, 2018). Along with these, Turkish readability was also checked using the Ateşman (1997) and Çetinkaya-Uzun (Çetinkaya, 2010) formulas.

After the operations were performed, the opinions of three Turkish field experts were consulted regarding the changes that might occur in the deep structures of the texts. The experts were presented first with the authentic and then the modified versions, and they were asked about the preservation status of the subject, main idea, character traits, and aesthetic value in these texts. Thus, it was aimed to preserve elements such as the main subject, main idea, character traits of individuals, plot, time, and setting in the texts. In the modifications made, efforts were made to preserve the aesthetic value by intensively performing sentence splitting operations and simplification with equivalent words. As a result of the feedback received, the text modification process was completed. Indeed, as discussed in the literature, it has been frequently stated that modified, intervened texts may undergo a "loss of aesthetic value" (Carroli, 2008; Gilmore, 2007; Young, 1999). Consulting expert opinion in order not to compromise the specific value of the text while making modifications gains importance here in terms of ensuring that the relevant intervention does not create a loss of value in the text (Uygur, 2024). On this occasion, the authentic value, thematic structure, and emotional depth of the text were attempted to be preserved with the help of expert opinion and corpus-based tools. Subsequently, the stage of writing reading comprehension questions for both texts was initiated. At this stage, open-ended questions were written, and the questions were given their final forms with expert opinions. Detailed information regarding modification and reading comprehension questions is presented below.

Modifications Regarding the Sümbül Kokusu Text: "Sümbül Kokusu" (Scent of Hyacinth), which can be considered a historical text, was subjected to a multi-layered modification process specific to the text type so that 7th-grade students could more easily access the text's emotional depth and historical context. In this process, the simplification strategy was predominantly utilized. In determining words not frequently used in contemporary Turkish, objective sources such as A Frequency Dictionary of Turkish (Aksan et al., 2016) and A Frequency Dictionary of Contemporary Turkish (Tezcan-Aksu & Adalı, 2018) were used alongside the intuitive evaluations of the researchers. These lexical simplifications covered expressions such as muzmahil olmak (to collapse), metrûk (abandoned), civanmert (noble-hearted), me'yûs (hopeless), and şehbenderhane (consulate). These were replaced with current and understandable equivalents in a way that would preserve the literary atmosphere of the text. In this process, in order to preserve the literary texture of the text, the modification of words whose meanings students were thought to be able to deduce from the context was avoided. In addition to these arrangements at the lexical level, long and descriptive sentences filled with nested clauses that slowed down the narrative flow of the text were split into shorter and dynamic sentences to make the plot more followable. Another important intervention specific to this text was the application of the shortening strategy. Some long descriptive passages aimed solely at creating atmosphere, which did not directly contribute to the main conflict of the text and the inner world of the characters, were removed from the text to keep the reader's attention on the main plot. For example, the paragraph starting with "Here is a summer evening, Kalender's cheeks..." in the authentic text, which describes a summer evening, was removed from the text because it did not directly contribute to the progression of the story. All these and similar interventions aimed to ensure that the target audience understands the spirit of the National Resistance and the internal conflict experienced by the heroes in the authentic text more easily and establishes a stronger bond with the text.

To conduct a preliminary evaluation of these interventions made, both the authentic and modified texts were read by the first study group consisting of 22 students. According to what these students stated, the words and phrases not understood in the authentic text were in the following contents: muzmahil (13), metruk (14), şehvani (13), revolver (13), pazı (13), dimağında (13), tayyarelerine (12), tahtelbahirlerine (12), mukabele (12), ma'mur (12), ma'mulatı (12), darü'l felasifelerinden (12), te'cil (11), tahsil (11), tabiiyyat (11), nurani (11), me'yus (11), kaşanelerinden (11), işportalar (11), civanmerttir (11), cerideyi (11), revolverini (10), mühimmat (10), muzlim (10), ma'muresi (10), istinatgahı (10), gamlı (10), dretnotlarına (10), şebboylar (9), rayihalar (9), matemli (9)... A total of 148 words/phrases were indicated in the list continuing in this manner. Students marked on the text the sections where these words appeared as not understood. It was reported by students on the text that sections with many unknown words were also not understood. In the modified text, 56 words with unknown meanings were reported. The most frequently repeating examples from these sections are as follows: pazı (11), loş (8), viranesi (7), takatsizdi (7), müracaat (7), şebboylar (6), ud (6), simalarıyla (6), gamlı (5)...

When the authentic text was examined according to Ateşman (1997) and Çetinkaya-Uzun (Çetinkaya, 2010) readability formulas, the results in Table 1 were obtained. Regarding the word and sentence counting directive, Çetinkaya (2010) was taken into account. According to the obtained results, it is observed that the Ateşman (1997) readability score increased by approximately two points. The existence of sections directly removed from the text, the frequent inclusion of word modification in the text, and the fact that sentence shortening was not done excessively may have been effective in this situation. The primary goal in the interventions made was that the artistic expression of the text should not be damaged. Therefore, processes of oversimplifying the sentence structure that could affect the expression were resorted to less.

Table 1. Readability Results of the Narrative Text

Text	Sentence Count	Word Count	Syllable Count	Character Count	Ateşman	Çetinkaya-Uzun
Authentic	116	882	2402	5896	69.57 (Medium)	40.67 (Educational Reading – 8th and 9th Grade)
Modified	113	762	2083	5128	71.40 (Easy)	41.24 (Educational Reading – 8th and 9th Grade)

The comparison of the modified text created after these stages and the authentic version is as follows (strikethrough indicates removed or simplified sections, and underlined sections indicate additions made):

Sümbül Kokusu

Bugün pazardı. Budapeşte ~~Darülfünunu Tabiiyyat şubesinde tahsil etmekte~~Üniversitesinde Doğa Bilimleri alanında okumakta olan Hüseyin Arif, ~~Macar payitahtının loş,~~Macaristan başkentinin rutubetli, güneşsiz ve loş sokaklarının birinde, bu milletin talihi gibi, gamlı, ~~metrûk~~boşaltılmış bir apartmanında ~~fakirane~~yaşamaktaydı. Yoksul odasının kenarında, tahta bir masanın üstünde duran bir gün evvelki ~~ceridevi~~gazeteyi aldı. Tekrar okumaya başladı. Çanakkale'ye dair ~~Dersaadet'ten~~İstanbul'dan, Berlin'den, Londra'dan gelen telgrafları ~~süzdü~~inceledi! Karadan ve denizden hücum devam ~~ediyor~~,ediyordu. Boğaz'a döktüğümüz ~~torpiller toplanıyor,~~mayınlar toplanıyordu. Kumkale'ye, Seddülbahir'e, Anafarta'ya çıkan İngiliz, Fransız askerleri ~~ilerliyorlar…~~ilerliyorlardı…

Her satır bir hançer, her nokta bir kurşun gibi beynine saplanıyordu. ~~Muzmahil idi~~Çökmüştü. Yüreğini kaplayan ateşli bir acı göğsünü yakıyordu. Romanya, ülkesinden mühimmat geçirtmiyordu. Bizim de silah ve cephane tezgâhlarımız ~~kâfi~~yeterli değildi. İstanbul'un ~~müdafaasından~~savunmasından ümidini kesmişti. Hâlbuki İngilizler, Fransızlar, cihanın bu iki ~~müterakkî~~gelişmiş milleti yerden, gökten, denizden, yıldırımlarıyla, volkanlarıyla, cehennemleriyle, kapımızın önüne gelmişlerdi. Onların ~~dretnotlarına, tayyarelerine, tahtelbahirlerine mukâbele edecek nemiz~~zırhlı gemilerine, uçaklarına, denizaltılarına karşılık verecek neyimiz vardı? Bir göğüs, bir pazı... İşte bu kadar!

Hay Allah'ım! Bu göğüs çökecek, bu pazı kıvrılacak.

Şimdi vatanı, İstanbul; bütün ~~camileriyle, saraylarıyla~~camileri, sarayları, mavi göğü, mavi ~~deniziyle, saz benizli,~~denizi, solgun yüzlü narin ~~kadınlarıyla~~kadınları, ince uzun boylu sinirli ~~gençleriyle~~gençleri, ağır ve ~~me'yûs~~ümitsiz yürüyüşlü ihtiyarlarıyla gözünün önüne ~~geliyorlar~~geliyor…

İslâm'ın gözü, ~~Türkün~~Türk'ün kalbi olan ~~bu,~~ renk ve nur durağı ~~memleket~~İstanbul pek temiz, pek ~~ma'mûr~~bakımlı, pek güzeldi. Onun yıkık duvarları, Avrupa'nın dargın sisler, durgun isler altında, kaba, kirli, kara, ~~matemli kâşânelerinden~~yaslı köşklerinden daha güler yüzlüdür~~…~~. Onun çarpık kavuklu, yan fesli harap mezarlıkları, buraların ~~Dârü'l-felâsifelerinden~~felsefe fakültelerinden, kütüphanelerinden daha manalı, daha düşündürücüdür~~…~~. Oranın hamalları, fakirleri buranın lortlarından, milyonerlerinden daha asil, daha ~~civanmerttir…~~yüce gönüllüdür. Buranın düzgün, kara sokaklarından oranın beyaz, mavi kaldırımlı eğri yolları, daha ~~nûrânî~~ışıklı, daha neşelidir~~…~~.

Avrupa'nın bir şehir büyüklüğündeki fabrikalarından, içlerinde bir ~~hünerverle~~ustayla bir çırak çalışan küçük ~~dükkânların ma'mûlâtı~~dükkânlarının ürünleri daha sanatlı, daha kıymetlidir…

Kalbinin en derin, en ateşli köşesinden bir çığlık kopardı:

– Allah! Ey İslâm'ın Allah'ı!.. Düşman, vatanımı çiğnemesin, çiğnetme!

Bulutlu düşünceler ~~dimağını~~zihnini kapladığı sırada uzun siyah kirpikleri arasından iki şimşek çaktı. Damla damla ağladı. Bu bir rahmetti. Gözlerini sıktı. Dar ve ~~muzlim~~karanlık odacığında gezinmeye başladı… ~~Sonra masasının gözünü çekti. Ufak, paslı bir revolver çıkardı. Fişeklerini yokladı. Ninesinin, ablasının, sevdiklerinin hayallerini gözleriyle öptü. Kâğıt ve kalem aldı…~~ Takatsizdi. Avluya ~~nâzır~~bakan pencereyi açtı. Bir parça hava istiyordu.

Pencere önünde duran ve dört gün evvel ev sahibesinin hediye ettiği küçük saksıdaki sümbüle doğru eğildi. Onu kokladı… İrkildi… Çekildi. Düşündü… Ağlayarak tekrar saksının üstüne kapandı… Kaldı… Sağ avucuyla çiçeği yüzüne, gözüne çekti, sürdü. Yine kokladı, kokladı, kokladı. ~~Dimağında~~Zihninde bir baygınlık, gönlünde bir aşk, bir secde ~~istiğrakile~~ kendinden geçme isteği vardı… Kendisini kaybetti.

Bu sırada kapı vurulmuştu. Dudakları çiçeğin kıvrımları arasında olduğu hâlde ~~mecalsiz~~hâlsizce "Gel!" dedi. Karşısında Mehmet Siyavuş'u buldu. Hemen saksıyı kavradı.

– Şunu kokla, kokla Mehmet, dedi.

Mehmet Siyavuş arkadaşının perişanlığından, heyecanından ürkmüştü. Yavaşça içini çekti.

– Hayır, iyi kokla! Derin kokla! Gözlerini kapayarak kokla! Koklarken gözlerinin önüne ne geliyor? Neresi geliyor? Söyle, Allah aşkına bütün ruhunla, bütün nefesinle kokla!..

Arkadaşının yüzüne doğru çiçeği tutuyor, itiyordu. Mehmet Siyavuş:

– İstanbul kokusu!

– Değil mi? Değil mi? Fakat neden böyle?

– Hani mart içinde, nisan, mayısta Köprübaşı'nda, sokak köşelerinde geniş işportalar, kola takılan ince uzun sepetler içinde laleler, zerrinler, şebboylar, menekşeler. "~~Bahariye~~İlkbahar kokuları!" diye bağıran kara yağız~~, bıçkın kıyafetli~~ satıcıların önünde demet demet saçılan bu ~~râyihalar~~güzel kokular… Beşiktaş'ın, Eyüp'ün fulya tarlaları kokuları… ~~Ruhanî~~Ruhani bir medeniyet, ~~kutsi~~kutsal bir nezaket kokuları…

Ah! Vatan!

"~~Misk~~Mis gibi kokusu canlarda tüter."

Güneşi böyle, göğü böyle kokar değil mi? Viranesi böyle, ~~ma'mûresi~~bakımlı yeri böyle kokar. Sarayı böyle, kulübesi böyle kokar değil mi? Fakat bu mübarek bahçe elimizden gidiyor, İstanbul'u kaybediyoruz.

– …

Şimdi korkunç bir ~~sükût~~sessizlik… İki genç boyunları bükülmüş, kalpler durmuş, gözleri fırlamış düşünüyorlardı…

– Evet, nasıl karşı duracağız?

– Nasıl karşı duracağız?

– Bu sefer mutlaka… Yok, söylemeye dilim varmıyor.

– Oooof!

~~– Allah aşkına şu sümbülü bir daha kokla!~~

~~İşte bir yaz akşamı, Kalender'in yanakları, saçları, elleri okşayan yumuşak, tatlı, ılık, şehvanî rüzgârı ki ruhu mavi nurlara sarar… İşte küpeştesi mavi ipek örtülü iki çifte narin bir kayığın içinde ak bulutlara bürünmüş iki yıldız, iki yaşmaklı taze hanım, Şair Nedim'in iki berceste mısraı; iki demet sümbül… İşte, yıldızlı bir gece, Anadolu sahilinin arka koruları arasından gelen bülbüllerin elhânına, görünmez bir sandaldan yükselen bir "Hey hey" ahengi bürünüyor, sanırsın ki yıldızlar nur teganni ediyorlar ve çiçekler rayiha ötüyorlar.~~

– Şu sümbülü ver, bir daha koklayayım. İşte, Babıali Caddesi'nden geçen ~~iki üç efendi ki~~ sarımtırak simalarıyla iki üç efendi, zarafeti dünyanın hiçbir tarafında inkâr ~~olunamayan~~edilemeyen asil ve ~~vakur inhinâlarıyla, aşinalarıyla temenna ediyorlar~~ağırbaşlı eğilmeleriyle, tanıdıklarıyla selamlaşıyorlar… İşte Bayezit Camii'nde bir teravih namazı ki bütün bir milletin saf ruhu secde hâlinde, Allah'ın huzuruna serilmiş… İşte, bir eski kafesin arkasında ~~şarkın~~doğunun bütün ~~enîn-i şiirini okuyan~~şiirlerini inleten bir ud ~~sedası~~sesi! İşte Bolayır'da Namık Kemal'in mezarı! Azizim, camileri, medreseleriyle, ~~imaretleri,~~ saraylarıyla bütün bir medeniyetin, üç yüz ~~milyon~~milyonluk İslâm'a ait bir medeniyetin, bir ~~mevcudiyetin istinatgâhı~~varlığın dayanak yeri, Türk'ün, ~~Osmanlının güzelliklerin, büyüklüklerin payitahtı~~Osmanlı'nın güzelliklerinin, büyüklüklerinin başkenti gidiyor.

– Gidiyor ha!

~~Hüseyin Arif şimdi fırlamış "Yaşamak alçaklıktır" diyerek revolverini tekrar kavramıştı. Siyavuş üstüne atıldı. "Hayır, sanırım ki böyle odada ölmek alçaklıktır" dedi.~~

– Ne yapabiliriz?

– Çanakkale'de ölebiliriz.

– Bundan ne olur?

– Analarımızdan evvel kendimizi çiğnetmiş oluruz. Ben gidiyorum.

Hüseyin Arif:

– Ben de geliyorum Mehmet.

Birbirlerinin ellerini sıktılar.

O akşam lokantada ~~rast geldikleri~~rastladıkları bir gazeteci bunlara demişti ki Çanakkale ~~müdafaası~~savunması yalnız ~~hilafetin~~Türkiye'nin istikrarını değil, Almanya ve Avusturya-Macaristan'ın da ~~bekasını temin edecek ve~~kalıcılığını sağlayacaktır. Bu, Rusya'nın ~~izmihlalini~~yok oluşunu hazırlayacaktır. Ve tarih, Türkiye'nin ~~ehemmiyetini~~önemini o zaman anlayacaktır.

İki gün içinde eşyalarını sattılar. Üçüncü gün pasaportlarını vize ettirmek için ~~şehbenderhaneye~~konsolosluğa müracaat ~~ettikleri~~ettiler. Bu sırada "~~Talebenin~~Öğrencilerin askerlikleri ~~te'cîl edildi"~~ertelendi." diyen ~~kâtibin~~memurun duygusuz gözlerine baktılar ~~ve~~. Ardından ceketlerinin ~~iliklerine~~yakalarına taktıkları bir sümbül çiçeğini derin derin koklayarak ve önlerine bakarak usulca, "Biz gönüllü gidiyoruz!" dediler.

Ahmet Hikmet Müftüoğlu

Çağlayanlar, 3 Şubat 1334

(Değiştirilmiştir)

Modifications Regarding the Dil Nedir? Text: The text "Dil Nedir?" (What is Language?), possessing academic and abstract content, was subjected to a unique and systematic modification process to increase the conceptual clarity and logical consistency required by the informative text type. In the lexical simplification stage of the text, words not frequently used in contemporary Turkish were determined using A Frequency Dictionary of Turkish (Aksan et al., 2016) and A Frequency Dictionary of Contemporary Turkish (Tezcan-Aksu & Adalı, 2018), as well as based on the intuitive evaluations of the researchers. Accordingly, academic terms such as içtimai (social), müessese (institution), vasıta

(tool/means), kaide (rule), and hüviyet (identity) were updated with more common scientific language usages while preserving semantic accuracy. Beyond lexical simplification, long, complex, and nested sentence structures, which were the most fundamental problematic aspect of the text, were analyzed. The fact that the average sentence length of the authentic text was 11.8 words made this syntactic intervention mandatory. These structures, harboring philosophical arguments, were divided into simpler and more understandable sentences, each expressing a single logical step. The most critical strategy specific to this text was elaboration (that is, semantic expansion) applied in sections that were important within the text but deemed insufficient to be explained at the middle school level. The most prominent example of this is the concretization of the expression "language is a system of secret agreements," which was thought to create uncertainty for the reader, with the sentence "That is, its beginning is not determined and how it functions is uncertain."

To evaluate the effectiveness of these applied strategies and the increase in the text's accessibility, the authentic and modified texts were presented to a pilot group of 22 students. According to what the students stated, the words and phrases not understood in the authentic text were as follows: içtimai (22), tabii (21), müessesedir (20), vasıta (17), mefhumları (11), kaideleridir (10), temayyüllerinden (9), meçhulümüzdür (9), hüviyete (9), fertlerin (9), cemiyetin (9), mahiyeti (8), hususiyetlerine (8)... A total of 102 words/phrases were indicated in the list continuing in this manner. Students marked on the text the sections where these words appeared as not understood. It was reported in writing by students on the text that sentences containing many words with unknown meanings and long sentences were also not understood. In the modified text, 50 words or phrases with unknown meanings were reported. Examples from these sections are as follows: zira (7), muhafaza (4), kavmi (4), olageldiğine (3), istisnalar (3), ilkelerdir (3), fertleri (3), ...

When the authentic text was examined according to Ateşman (1997) and Çetinkaya-Uzun (Çetinkaya, 2010) readability formulas, the results in Table 2 were obtained. Regarding the word and sentence counting directive, Çetinkaya (2010) was taken into account. According to the obtained results, it is observed that the Ateşman (1997) readability score increased by seven points, and the Çetinkaya-Uzun (Çetinkaya, 2010) readability score increased by three points. Thus, it can be interpreted that the text's readability level became easier.

Table 2. Readability Results of the Informative Text

Text	Sentence Count	Word Count	Syllable Count	Character Count	Ateşman	Çetinkaya-Uzun
Authentic	70	827	2197	5332	61.26 (Medium)	38.31 (Educational Reading – 8th and 9th Grade)
Modified	79	795	2063	5012	68.31 (Medium)	41.62 (Educational Reading – 8th and 9th Grade)

The comparison of the modified text created after these stages and the authentic version is as follows (strikethrough indicates removed or simplified sections, and underlined sections indicate additions made):

Dil Nedir?

Dil, insanlar arasında anlaşmayı sağlayan doğal~~tabiî~~ bir araçtır. O~~vasıta~~, kendisine özgü~~mahsus~~ kanunları olan ve ancak bu kanunlar çerçevesinde gelişen canlı bir varlıktır. Temeli~~varlık, temeli~~ bilinmeyen zamanlarda atılmış ~~bir~~ gizli anlaşmalar sistemidir. Yani başlangıcı belirli değildir ve nasıl işlediği belirsizdir. Ayrıca~~antlaşmalar sistemi~~, seslerden örülü toplumsal~~örülmüş içtimaî~~ bir yapıdır~~müessesedir~~.

Doğal Bir Araç ~~Tabiî bir vasıta~~

Dil, insanlar arasında anlaşmayı sağlayan doğal~~tabiî~~ bir araçtır~~vasıtadır~~. İnsanlar; duygularını, düşüncelerini, fikirlerini~~, hükümlerini~~ birbirlerine aktarmak, isteklerini~~nakletmek, meramlarını~~

birbirlerine anlatmak için dil denen araca başvururlar.~~denilen vasıtaya baş vururlar.~~ Fakat dil insanların kullandığı herhangi bir araca~~vasıtaya~~ benzemez. Onun araçlığı~~vasıtalığı~~ sadece anlaşmayı sağlaması bakımındadır. Bireyler~~temin etmesi bakımındandır. Fertler~~ ve nesiller arasında anlaşma aracı~~vasıtası~~ olarak iş görür. Fakat bu işi görürken daima bağımsız~~müstakil~~ bir kimliği vardır.~~hüviyete sahiptir.~~ İnsanlar ona istediği~~istedikleri~~ gibi hükmedemezler. Yani onu kontrol edemezler. Onu olduğu gibi kabul etmeye~~etmeğe~~, onu bir araç~~vasıta~~ olarak kullanırken onun özelliklerine~~hususiyetlerine~~ dikkat etmeye~~etmeğe~~, onun doğasına uymaya~~hususiyetlerine uymağa~~ mecburdurlar. Çünkü dil yapay~~sun'î~~ bir araç, maddi~~vasıta, maddî~~ bir araç~~vasıta~~, bir âlet değildir. O doğal~~tabiî~~ bir araçtır. Araç görevi~~vasıtadır. Vasıta vazifesi~~ görür~~,~~ fakat doğal~~tabiî~~ bir varlığı vardır. Bu~~Dil bu~~ bakımdan dil, canlı bir araca~~vasıtaya~~ benzer.

Mesela~~Meselâ~~ at da bir araçtır~~vasıtadır~~, otomobil de bir araçtır. İnsan otomobili~~vasıtadır. Fakat insan otomobile~~ istediği şekilde yönlendirebilir fakat atın doğasına~~hükmedebilir, at karşısında ise ancak onun tabiatına~~ uygun hareket etmek zorundadır. Otomobile istediği şekli verir, onun biçimini istediği şekle sokar, onu istediği gibi kullanır, isterse uçuruma götürebilir.~~sevk edebilir.~~ Fakat atın biçimini değiştiremez, onu istediği gibi kullanamaz, istediği yere götüremez. Ata ne ederseniz edin,~~sevk edemez. Başını kesseniz ata~~ korktuğu yerde bir adım attıramazsınız. İşte dilin araçlığı da böyledir~~vasıtalığı böyle bir vasıtalıktır~~, atın araçlığı~~vasıtalığı~~ gibidir. Anlaşmayı sağlamak bakımından bir araç gibi iş görür fakat doğal bir varlığa sahiptir.

Canlı Bir Varlık~~bir varlık~~

Doğal~~Tabiî~~ bir araç~~varlık~~ olan dilin kendisine özgü~~mahsus~~ birtakım kanunları vardır. Bunlar dil kurallarıdır~~kaideleridir~~. Dil kuralları,~~kaideleri~~ dilin yapısına hâkim olan, dilin bünyesinden ve eğilimlerinden doğan bazı ilkelerdir.~~yapısından ve temayüllerinden doğmuş bulunan birtakım prensiplerdir.~~ Bunlar dille birlikte var~~mevcut~~ olup onun yapısının özelliklerini gösterir.~~hususiyetlerini ifade ederler, temayüllerinin istikametlerini gösterirler.~~ Dil, canlı bir varlıktır.

Zaman zaman birtakım değişiklikler geçirebilir. Kendi yapısından~~, kendi bünyesinden~~ doğan çeşitli sebeplerle de bazı gelişmeler gösterir. Bu gelişmeler ve değişiklikler ~~ve gelişmeler~~ ona~~,~~ uzun tarihi boyunca~~,~~ daima gelişen~~serpilen~~ ve zaman içinde akıp gelen bir görünüm~~manzara~~ verirler.

Bu yüzden dilin tarihinde birtakım aşamalar~~merhaleler~~, birtakım gelişme evreleri~~safhaları~~ göze çarpar. Fakat bütün bu değişiklikler ve gelişmeler dil kuralları~~kaideleri~~ çerçevesinde ortaya çıkar.~~cereyan ederler.~~ Dile yeni kelimeler kazandırmak için dışarıdan yapılacak müdahalelerin de ~~daima~~ bu kurallar~~kaideler~~ çerçevesinde olması gerekir. Kendi kanunlarına aykırı zorlamaları dil hiçbir zaman benimsemez. Canlı bir varlık olarak yapısı bireylerin~~, fertlerin~~ ve toplumların ona karışmasına hiç~~cemiyetlerin istedikleri şekilde~~ karışmalarına müsait~~, uygun~~ değildir. Onun bireylere~~fertlere~~ ve toplumlara bağlı~~cemiyetlere tâbi~~ olmayan bir düzeni~~nizamı~~ vardır. Bu düzeni~~nizamı~~ meydana getiren şey, kendi kanunları ve kurallarıdır.~~, kendi kaideleridir.~~ Bu kurallar dışındaki~~kaideler dışına çıkacak~~ bir müdahale dile hiçbir şey kazandırmaz. Dil ancak kendi bünyesine uygun normal bir müdahaleyi kabul eder. Normal bir müdahale ise sadece dilin doğal gelişim~~tabiî gelişme~~ yolunu açık tutmaktır. Yani yalnızca~~dışarıdan dile, ancak,~~ dilin doğal gelişimini~~tabiî gelişmesini~~ önleyen bir durum varsa dile dışarıdan~~,~~ müdahale edilmelidir. Zira bazen dilin bünyesini saran yabancı unsurlar yabancı~~, zararlı~~ otlar gibi gelişimine zararlı olur. Böyle durumlarda dilin doğal gelişme yolunu açık tutmak için, dile yardım etmek için müdahale etmek lazımdır. Böyle bir yardım ise ancak dile kendi kuralları içinde kalarak yaklaşmak şartı ile faydalı olabilir. Çünkü dil kendi kanunları, kendi kuralları içinde gelişen canlı bir varlıktır.~~, dilin tabiî gelişmesine engel olurlar. Böyle durumlarda dilin tabiî gelişme yolunu açık tutmak için yabancı unsurları temizlemek üzere dile dışarıdan yardım etmek mümkündür ve lâzımdır. Böyle bir yardım ise ancak dile kendi kaideleri içinde kalarak yaklaşmak şartı ile faydalı olabilir. Çünkü dil kendi kanunları, kendi kaideleri içinde gelişen canlı bir varlıktır.~~

Gizli Anlaşmalar Sistemiantlaşmalar sistemi

Dil bir gizli anlaşmalar sistemidir. Canlı ve cansız varlıkları, kavramlarımefhumları, hareketleri karşılayan kelimeler ve oüzerinde, kelimelerin nasıl sıralandığı üzerinde bir toplumun tüm birbirleri ile münasebetleri ve fikirleri anlatmak için yapılan kelime sırası üzerinde bir cemiyetin, bir kavmin, bir milletin bütün fertleri gizli anlaşmalar, gizli sözleşmeler yapmış durumdadır.durumdadırlar. Bu yollasuretle bir toplumun tüm bireylericemiyetin bütün fertleri bir varlığı hep aynı kelimeyleayni kelime ile karşılarlar. MeselaMeselâ bütün Türkler bildiğimiz sert şeyecisme taş, suya su, ışığa ışık demek için âdeta sözleşmiştir.sözleşmişlerdir. Bu sözleşmeyi, bu gizli anlaşmayı her toplumcemiyet, her kavim ayrı yarıbir şekilde yapmış, böylece her milletinkavmin ayrı bir dili olmuştur. AynıAyni bir varlığa Türkler taş, Farslar seng, Araplar hacer demiştir. Kelimelerin ilişkileridemişlerdir. Ayni duygu ve düşünceleri anlatmak için kelimelerin münasebeti ve sıralanışı da milletlere göre farklıdır.bu kavimlerde başka başkadır. Çünkü her toplumun kendine özgükavmin ayrı bir gizli anlaşmalar sistemi vardır. BuBir kavmin bütün fertleri arasında mevcut olan bu gizli anlaşma ve sözleşmelerin temeli bilinmeyen zamanlarda atılmıştır. Dil insanla birlikte var olageldiğineola geldiğine göre bu gizli anlaşmaların kökleri ilk insanlara kadar gider.

Yalnız, bu gizli anlaşmaların doğuşu ve niteliğimahiyeti bilinmemektedir. Varlıkların ve hareketlerin sözlerle karşılandığı gibi bu karşılayışın ayrı kavimlere göre farklı oluşunun da sebeplerini ve niteliğini bilmiyoruz. Bu konudaki bir görüşe göre dilin sebepleri ve mahiyeti meçhulümüzdür. Bu hususta mevcut kanaat dillerin doğuşunda doğadaki seslerin insanlar tarafından taklit edilmesi çok önemlidir.tabiattaki sesleri taklidin mühim bir yeri olduğu merkezindedir. Bugün her dilde ses taklitlerindentaklidinden doğdukları açık olan bazı kelimeler de vardırmevcuttur. Fakat istisnalar hariç sözcüklerin çoğunun böyle bir taklitbirkaç istisna dışında büyük kelime kütlelerinin herhangi bir ses taklidi izi taşımadığıtaşımadıkları da ortadadırmuhakkaktır.

Toplumsal ve Millî Bir KurumDil içtimaî müessesedir

Dil toplumsal bir kurumdur. BireylerinFertlerin üstünde, bütün toplumunbir cemiyetin malı olan ve bütün bir toplumucemiyeti içine alan kuvvetli bir kurumdur. Toplumlarınmüessesedir. Cemiyetlerin en büyük dayanağı dildir. Dil, bir toplumuBir cemiyeti ayakta tutan, bir toplumuncemiyetin varlığını sağlayan, devam ettiren, bir toplumacemiyette sarsılmaz bir birlik yaratan kurumdur. Bu bakımdan müessese olarak dilin oynadığı rol çok büyüktür. Bu bakımdan dil, milleti oluşturanteşkil eden unsurların başında gelir.

Bir milleti, bir kavmi bazen tek başına ayakta tutar, millî benliği muhafaza ederek, onu yok olmaktan, eriyip başkalaşmaktan kurtarır. Demek ki dil bir milletin en büyük millî kurumudur.müessesesidir. Bu toplumsaliçtimaî ve millî kurumunmüessesenin malzemesi ise seslerdir, yapısı seslerden örülmüştür. Sesler yan yana gelerek kelimeleri ve kelime dizilerini meydana getirir. O hâlde dil, seslerden yapılmış bir bütünbina, seslerden kurulmuş bir yapı, büyük bir sesler sistemi, seslerden örülmüş toplumsaliçtimaî bir kurumdur.müessesedir.

Muharrem ERGİN

(Değiştirilmiştir)

Comprehension Questions

Since the goal was for the comprehension questions to cover both the modified and authentic texts, the preparation of questions for the informative and narrative texts began after the text modification process. Thus, a 10-item pool was prepared by the researchers for both texts (informative and narrative). Three of these questions were written in accordance with higher-order thinking skills located at the analysis, evaluation, and creation steps of the taxonomy (Kanık-Uysal, 2022). Since studies in the literature indicate that multiple-choice questions often fail to go beyond the understanding and application steps (Kanık-Uysal, 2022), the questions in this study were structured as open-ended. A graded scoring key (rubric) was prepared for the evaluation of the questions. In this scoring key, lower-order questions were scored as 0, 1, 2; whereas questions regarding higher-order thinking skills were

scored as 0, 1, 2, 3. Following these procedures, opinions were obtained from three experts. Two of the experts are faculty members with PhDs in Turkish Education, and one in Measurement and Evaluation in Education. The experts evaluated the comprehension questions according to the criteria of (i) cognitive level according to Bloom's revised taxonomy, (ii) compliance with open-ended item writing principles in Turkish teaching, (iii) content validity, (iv) linguistic and structural clarity for students, and (v) suitability for scoring. These criteria are based on the open-ended item preparation and evaluation principles suggested in Özçelik (2016) and Brookhart (2010). In line with these expert opinions, the number of questions was reduced to seven for the informative text and eight for the narrative text. Two of the seven questions in the informative text are related to higher-order thinking skills, while the others are lower-order questions. The minimum score obtainable is 0, and the maximum score is 16. Of the eight questions in the narrative text, two are higher-order and six are lower-order questions. The minimum score obtainable is 0, and the maximum score is 18.

After these stages, a pilot application was conducted on 34 students in the second study group. In this application, it was examined whether the instructions of the questions were clear and understandable, the students' levels of correctly understanding the question, and whether the answers overlapped with the expected cognitive level. Structured criteria were determined for the evaluation of clarity. Accordingly, (i) conceptual clarity, (ii) suitability of the language used for the age and cognitive level, (iii) whether it contained ambiguity that could lead to polysemy, and (iv) whether it reflected the expected thinking level of the student were taken into account for each question. In addition to this evaluation, during the application, students were explicitly asked whether the questions were understood, whether their meanings were clear, and whether there were unknown or obscure words within the question. Furthermore, student answers obtained in the pilot application were classified to serve as examples for the graded scoring key, and the final scoring guide was created. In the scoring structure of the questions in the study, generally, 0 points means "empty or irrelevant answer," 1 point means "partially correct but incomplete or superficial answer," 2 points means "complete and correct answer," and 3 points (for higher-order questions) means "analytical or inference-based, comprehensive, and correct answer." To concretize the application of the rubric, for instance, to the question "According to the text, what is the reason for Hüseyin Arif being in Budapest?" in the narrative text, a student who gives an incomplete answer like "For education" receives 1 point, while a student who gives a complete and text-based answer like "To receive education in the field of Natural Sciences at the university" receives 2 points. The full version of the scoring key containing detailed grading for all questions is presented in the Appendices.

The questions prepared for the informative text are as follows:

1. According to the text, why must humans behave in accordance with the nature of language?
2. What could be the purpose of the analogy made with the horse-automobile comparison in the text?
3. Which features are emphasized in the text by stating that language is a living entity?
4. In your opinion, what could be the reasons for different nations expressing the same concepts with words different from other nations?
5. According to the text, how should interventions on language be, and when can they be made?
6. Explain the expression "Language is a system of secret agreements" based on the text.
7. Do you agree with the idea that language has a protective effect for a nation? Why? Explain with your reasons.

The reading comprehension questions prepared for the narrative text are as follows:

1. According to the text, what is the reason for Hüseyin Arif being in Budapest?
2. Give two examples from the text regarding the attitudes of Hüseyin Arif and Mehmet Siyavuş showing their patriotism.
3. In which aspect are the British-French and the Turks on the battlefield compared in the text?
4. About which subject is Arif hopeless in the first part of the text?
5. What is the smelling of the hyacinth flower by the characters a symbol of?
6. What are the dominant emotions in the text?
7. What is the most comprehensive judgment that can be deduced from the text?
8. Do you find it right for the characters to leave Budapest and return to defend their country? Why?

Data Collection

The texts and reading comprehension tests prepared for data collection were administered to students in a face-to-face classroom environment. In all applications, the same instructions were given to the students; points such as which text they would read, how they would answer, and whether there was a time limit were explained verbally in a standard manner. Students were asked to read silently, answer the questions carefully after finishing the text, and request extra paper if needed. No guidance in the nature of directing, leading to the correct answer, or providing hints was provided to the students during the applications. No intervention was made other than timekeeping or test direction. All applications were carried out under the supervision of the researcher and the classroom Turkish teacher, without disrupting lesson hours. Additionally, 12-point font size and 1.5 line spacing were used in the printed materials in all applications.

The first study group was utilized in the text modification process, and applications were completed in a total of four lesson hours, being one lesson hour per week for four weeks. First, the modified and authentic texts were presented to the students in terms of words and sentences with unknown meanings. Students were asked to mark on the text the words and sentences they did not know the meaning of or had difficulty understanding. This stage aimed to provide student-focused data for text modification operations.

Data from the second study group were collected in two weeks. Authentic texts were presented to half of the group, and modified texts to the other half. The prepared draft reading comprehension questions were also answered at this stage. One text was presented to the students per week. No time restriction was presented to the students in each text application. However, it was observed that student readings and answers were completed in 60 minutes. At this stage, the answers given by the students and the comprehension process were used to evaluate the understandability and applicability of the questions.

The third study group is the group where experimental applications were carried out, and the data collection process lasted for three weeks. Before starting the application, the equivalence of the groups was checked with the reading comprehension test. The narrative text was administered in the first week of the application, and the informative text in the third week. No application was conducted in the second week. The application was carried out in two groups as the modified (experimental) and authentic (control) groups. Since the groups were selected from different schools, classes were not combined; four different classes were used. In the administration of both texts, students were instructed to read silently, not to use additional resources (dictionaries, etc.), and to answer open-ended questions with their own expressions. Instructions and application conditions were kept equal in all groups. Prior to the application, the equivalence of the groups was checked by administering a reading comprehension pre-test to both groups. Following the applications, post-test data were collected via reading comprehension questions.

Data Analysis

In the data analysis, the assumption of normality and other parametric conditions were considered; assumptions were checked for each test and reported in the findings section. Analysis of Covariance (ANCOVA) was used to examine the differences between groups while controlling for the effect of participants' pre-test scores. This method ensures that differences between groups are evaluated independently of baseline achievement levels. All analyses were conducted using the Jamovi (The Jamovi Project, 2023) open-access statistical software. Inter-rater reliability was calculated using the *seolmatrix* module available in Jamovi (Gamer et al., 2019; Seol, 2024).

Three expert raters were used in scoring the reading comprehension questions. The agreement between these raters was examined using Fleiss' Kappa, Krippendorff's Alpha, and Kendall's *W* coefficients. Fleiss' Kappa coefficient examines the agreement between scores given by more than two raters (Fleiss, 1971), and Landis and Koch (1977) stated that a result between .21 - .40 indicates fair agreement. Krippendorff's Alpha coefficient examines the agreement between any number of raters, does not require a minimum sample size, is compatible with nominal, ordinal, interval, ratio, and other measurement tools, and indicates perfect agreement as its value approaches 1 (Krippendorff, 2011). Kendall's *W* statistic also indicates the agreement between more than two raters with values between 0-1 (approaching 1 indicates high consistency between raters) (Field, 2024). Inter-rater reliability for the narrative and informative texts is given in Table 3. Accordingly, it can be stated that the agreement between raters is at an acceptable level. Therefore, the average of the scores given by the three raters was used in comparisons between groups.

Table 3. Inter-Rater Reliability of Reading Comprehension Questions

Text	**N**	**Raters**	**Fleiss Kappa**			**Krippendorff Alpha**	**Kendall's W**		
Narrative	90	3	Kappa	*Z*	*p*	Alpha	*W*	X^2	*p*
			.433	24.6	<.001	.952	.976	261	<.001
Informative	91	3	Kappa	*Z*	*p*	Alpha	*W*	X^2	*p*
			.337	17.2	<.001	.845	.919	248	<.001

Narrative_Text: Fleiss Kappa 95% CI (*k*=1000) = .358, .515; Informative_Text: Fleiss Kappa 95% CI (*k*=1000) = .265, .419

Ethics Committee Approval

The research was conducted with the permission of the Ordu University Rectorship Social and Humanities Research Ethics Committee, dated 15.06.2022 and numbered 2022-156. Additionally, ethical procedures suggested by the American Educational Research Association [AERA] (2011) and the American Psychological Association [APA] (2020) were meticulously followed.

Findings

In this section, ANCOVA findings regarding the comparison of scores obtained from reading comprehension questions for the narrative text and the informative text are presented, respectively. Additionally, comparisons of reading comprehension scores are also presented as graphs.

Table 4 presents the comparison of reading comprehension scores obtained by the groups administered the modified and authentic texts.

Table 4. ANCOVA Results of Reading Comprehension Scores for the Sümbül Kokusu Text

Source	*SS*	*df*	*MS*	*F*	*p*	η^2_p
Reading Comprehension (Covariate)	37.1	1	37.1	3.22	.076	.036
Group (Mod.-Aut.)	95.0	1	95.0	8.25	.005	.087
Residuals	1000.8	87	11.5			

The assumption of normality was verified by Shapiro-Wilk (*S-W* = .993, *p* = .913) and Kolmogorov-Smirnov (*p* = .929) tests. The assumption of homogeneity of variances was verified by Levene's test, $F(1, 88) = .308$, $p = .580$. *SS* = Sum of Squares; *MS* = Mean Square; η^2_p = Partial eta squared.

The narrative text application was carried out with a total of 90 students, 42 in the modified text group and 48 in the authentic text group. Subsequently, reading comprehension questions related to the "Sümbül Kokusu" text were directed to both groups. Prior to the analysis, assumption checks were performed. Normality was verified by Shapiro-Wilk and Kolmogorov-Smirnov tests ($p > .05$), and homogeneity of variances was ensured by Levene's test ($p = .580$). ANCOVA results showed that after controlling for pre-test scores, there was a statistically significant difference between the groups in terms of reading comprehension scores ($F(1, 87) = 8.25$, $p = .005$, $\eta^2_p = .087$). This reveals that the modified text group ($M = 11.2$, $SD = 3.54$) obtained significantly higher scores than the authentic text group ($M = 9.03$, $SD = 3.34$) (see Table 5). The calculated partial eta squared value ($\eta^2_p = .087$) indicates that this difference has a medium effect size (Cohen, 1988). The effect of pre-test scores, which was the covariate in the model, on post-test scores was not found to be statistically significant ($p = .076$).

Table 5. Descriptive Statistics for the Narrative Text

	Group	N	Missing	M	SD	Min.	Max.
Sümbül	Modified	42	1	11.19	3.54	2.00	17.3
Kokusu	Authentic	48	1	9.03	3.34	2.33	17.7

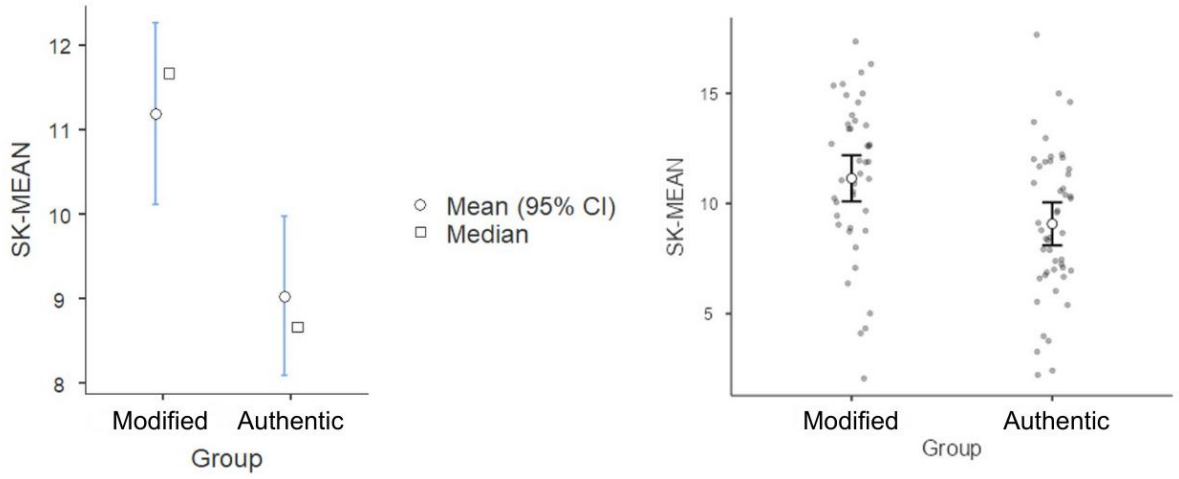


Figure 2. Comparison of Modified and Authentic Reading Comprehension Scores in the Narrative Text

In Figure 2, it is observed that the mean reading comprehension scores of the modified text group in the narrative text application are higher compared to the authentic text group. The graph shows the mean scores and 95% confidence intervals for both groups. The modified text group has a higher mean value, and its confidence interval is situated above that of the authentic text group.

The examination of reading comprehension scores for the modified and authentic informative texts is presented in Table 6.

Table 6. ANCOVA Results of Reading Comprehension Scores for the Dil Nedir Text

Source	*SS*	*df*	*MS*	*F*	*p*	η^2p
Reading Comprehension (Covariate)	26.2	1	26.22	3.77	.055	.041
Group (Mod.-Aut.)	100.7	1	100.73	14.48	<.001	.141
Residuals	612.2	88	6.96			

The assumption of normality was verified by Shapiro-Wilk (S-W = .989, p = .653) and Kolmogorov-Smirnov (p = .586) tests. The assumption of homogeneity of variances was verified by Levene's test, F(1, 89) = 2.04, p = .157. SS = Sum of Squares; MS = Mean Square; η^2_p = Partial eta squared.

The informative text application was carried out with a total of 91 students, 43 in the modified group and 48 in the authentic group. Reading comprehension questions regarding the text "Dil Nedir?" were directed to both groups. Prior to the analysis, assumption checks were performed. Normality was verified by Shapiro-Wilk and Kolmogorov-Smirnov tests ($p > .05$), and homogeneity of variances was ensured by Levene's test ($p = .157$). ANCOVA results showed that after controlling for pre-test scores, there was a statistically significant difference between the groups in terms of reading comprehension scores ($F(1, 88) = 14.48$, $p < .001$, $\eta^2_p = .141$). This reveals that the modified text group ($M = 7.89$, $SD = 3.05$) obtained significantly higher scores than the authentic text group ($M = 5.74$, $SD = 2.29$) (see Table 7). The calculated partial eta squared value ($\eta^2_p = .141$) indicates that this difference has a large effect size (Cohen, 1988). The effect of pre-test scores, which was the covariate in the model, on post-test scores is not statistically significant ($p = .055$).

Table 7. Descriptive Statistics for the Informative Text

	Group	N	Missing	M	SD	Min.	Max.
Dil Nedir	Modified	43	0	7.89	3.05	1.67	15.3
	Authentic	48	1	5.74	2.29	1.33	11.0

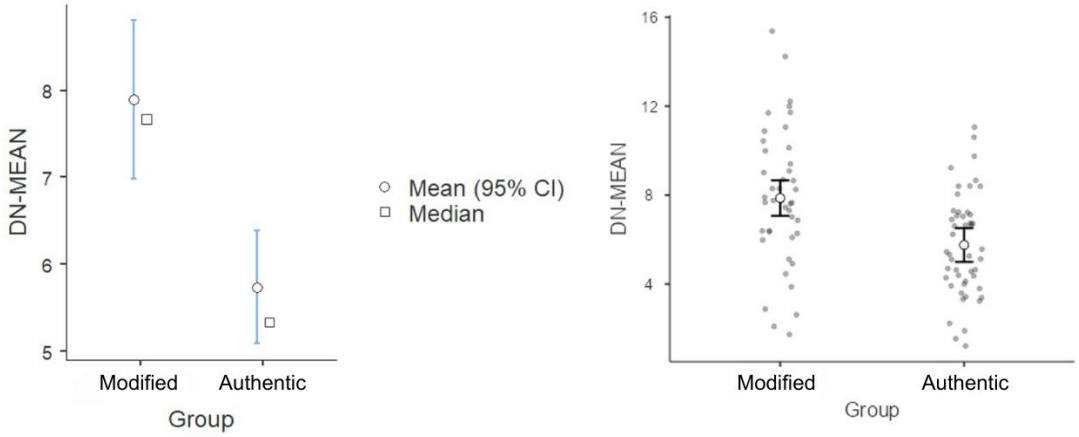


Figure 3. Comparison of Modified and Authentic Reading Comprehension Scores in the Informative Text

Additionally, the graph visually presenting the score difference is presented in Figure 3. When the graph is examined, it is observed that the modified text scores take higher values on average within the 95% confidence interval. The highest value the authentic text scores can take is lower than the lowest value the modified text scores can take.

Conclusion and Discussion

Within the scope of this research, authentic and modified versions of the same texts were presented to seventh-grade students, and reading comprehension questions prepared for these texts were administered. According to the findings obtained from the research, it was determined that students' comprehension levels of the modified text were significantly higher than their comprehension levels of the authentic text in both narrative and informative texts based on the quantitative results. In the measurement performed using open-ended questions, the mean score of the modified text group was calculated as 11.20 in the narrative text and 7.89 in the informative text. In contrast, the mean score of the authentic text group was 9.03 in the narrative text and 5.74 in the informative text. A difference of 2.17 points (M_mod. = 11.2, M_aut. = 9.03) was obtained in the narrative text application, and a difference of 2.15 points (M_mod. = 7.89, M_aut. = 5.74) in the informative text application. These differences correspond to approximately 12-13% of the total score in both tests. These general results demonstrate the effectiveness of text modification. However, the fundamental contribution of this study to the literature is presenting with quantitative evidence how the effect differentiates depending specifically on the text type, rather than repeating the general finding. The most original and central finding of the study is showing that the effect size of text modification in informative texts (large effect) is statistically much more distinct compared to narrative texts (medium effect). This differentiated effect can be explained by the difference in cognitive support systems that text types offer to the reader. Narrative texts possess natural narrative scaffolds such as plot, character development, and causality chains that help the reader construct meaning. It may be possible for the reader to follow the general meaning of the text by overcoming certain linguistic difficulties thanks to these scaffolds. In contrast, as stated in the literature, the fact that expository texts are more challenging than narrative texts (Dickens & Meisinger, 2017) may stem from these texts being conceptually more dense and linguistically more fragile. Meaning in such texts depends largely on the correct decoding of specific terms and complex sentence structures. Furthermore, since narrative scaffolds are fewer, linguistic barriers carry the potential to directly transform into a comprehension barrier. Therefore, the stronger effect of text modification on informative texts can be explained by this intervention directly targeting the most fundamental comprehension barrier. This finding offers an important theoretical perspective to the literature by suggesting that text type can be a critical factor in the success of the intervention.

In applications conducted with both narrative and informative texts, reading comprehension scores for modified texts are higher than those for authentic texts. The data are consistent with the results of studies in the literature. Applying elaborative modifications and improving coherence within the text content enhance the reading comprehension skills of native speakers (Beck et al., 1984). Bayraktar and Durukan (2020), in their application employing surface structure changes on both narrative and informative texts, determined that modified texts were understood better and stated that students reported positive views regarding these changes. The simplification of texts significantly improves reading fluency and comprehension, particularly for weak readers and those with weaker cognitive skills (Javourey-Drevet et al., 2022). It is also observed in the literature that text simplification is particularly conducted for weak readers and provides benefits. Text simplification may play an important role in improving reading comprehension for individuals experiencing difficulty in understanding informative texts, especially when the cognitive and linguistic needs of readers are taken into consideration (Arfé et al., 2018). It has been observed that texts simplified via the Open Book tool are more easily understood by both adults and adolescents with high-functioning autism (Cerga-Pashoja et al., 2019). Text simplification can have a positive effect on reading fluency and comprehension in children with reading difficulties (Gala et al., 2018). Along with these, it is stated that text difficulty level is also influential on reading comprehension. In primary school students, increased text difficulty is generally associated with decreased reading fluency and is negatively related to reading comprehension (Amendum et al., 2018). The reading proficiency of third-grade students is similarly affected by text complexity and length; comprehension in longer texts at the same level of complexity is lower compared to shorter texts (Mesmer & Hiebert, 2015). Text complexity negatively affects reading outcomes, particularly oral reading fluency and free recall, in middle school students (Spencer et al.,

2018). Within the scope of the research, reducing the difficulty level of the texts was also aimed. The detection of a difference in terms of reading comprehension in the modified texts supports these findings.

Simplification in vocabulary was frequently resorted to in the text modification applied in the research. In the pilot application conducted after modification, students also stated that words with unknown meanings were more frequent in the authentic text. The result of the research also supports that the simplification of words within the text can be beneficial in terms of reading comprehension. A meta-analysis of 37 studies investigating vocabulary interventions from kindergarten to 12th grade showed that such interventions improve comprehension, provide benefit especially to students with reading difficulties, and a correlation was observed between vocabulary and comprehension effects (Elleman et al., 2009). In the study conducted by Javourey-Drevet et al. (2022), simplified texts were read faster than authentic texts. In this study, we see that four of the students stated that the authentic text contained a large number of unknown words, and therefore they read the authentic text more slowly. Morphological knowledge significantly affects reading comprehension, contributing 2% to 9% of unique variance to reading comprehension (Foorman et al., 2012), and awareness regarding the structure and meaning of morphologically complex words plays a critical role in this process (Carlisle, 2000). As can be seen, vocabulary and morphological awareness related to words have been important factors in terms of reading comprehension in the research conducted. The cognitive mechanism underlying this vocabulary effect can be explained by the relationship between the reader's limited working memory capacity and the cognitive load created by the text. Reading is a complex action requiring both lower-level cognitive processes such as word recognition and higher-level cognitive processes such as understanding the whole text simultaneously. The existence of a direct relationship between reading skill and working memory capacity is also investigated (Daneman & Carpenter, 1980). When readers encounter a large number of words whose meanings they do not know, they may be compelled to allocate a significant portion of their limited mental resources to lower-level processing procedures such as decoding these words. This situation may leave fewer cognitive resources for higher-level comprehension processes such as constructing the general semantic structure (situation model) of the text and making inferences (Kintsch, 1998). At this point, Cognitive Load Theory offers an important explanatory framework. According to this theory, an effective instructional design must minimize extraneous cognitive load, which is not directly related to learning and leads to unnecessary mental effort (Sweller et al., 2019). When viewed from this perspective, the text modification in our study can be seen as a specific instructional intervention that reduces this extraneous load created by the effort to decode unknown words, particularly through lexical simplification. Therefore, the success of this intervention can be evaluated not only as eliminating unknown words but also as a cognitive support mechanism that helps the reader mobilize their mental resources for comprehension by using their limited working memory more efficiently (Baddeley, 2003). Furthermore, in research conducted on the reading habits of middle school students, students preferred the number of unknown words in the texts they read to be few rather than none (Aydın & Ayrancı, 2018). In the modified text as well, words whose meanings were not directly known were left in the text by preserving or highlighting the context. Thus, knowing most of the text vocabulary rather than all of it can also be characterized as effective on reading.

The effect of readability on the results obtained, as in this study, was previously determined in a study conducted by Durukan (2014) using different texts. In that study, a direct correlation was detected between the readability levels of the texts and comprehension rates. Furthermore, adolescents' reading comprehension is affected by both readability and cohesion, and texts that are easier to read and have high cohesion significantly increase adolescents' performance (Reed & Kershaw-Herrera, 2016). According to Bağcı-Dağdeviren (2024), for effective reading instruction, readers having the opportunity to read texts suitable for their levels can be effective in developing reading comprehension and a positive attitude toward reading. Text simplification helps students with low literacy levels, non-native English speakers, and those struggling with reading comprehension problems by increasing readability and understandability (Al-Thanyyan & Azmi, 2021). Li et al. (2005) determined in the research they conducted with Filipino high school students that texts subjected to intervention in the

form of simplification and elaboration increased the level of reading comprehension. Similarly, Long and Ross (1993) determined that Japanese university students reached a better level of comprehension thanks to L2 simplified texts. Additionally, in the study conducted by Yano et al. (1994), it was found that foreign language learning students who read linguistically simplified texts obtained significantly higher scores in comprehension compared to those who read authentic texts, but there was no significant difference between simplification and elaboration. On the other hand, Maxwell (2011) determined that text modification did not lead to a significant effect when primary school level English learners and informative texts were in question, and suggested that such studies should not be generalized to all grade levels. In the study of Uygur et al. (2025), it was also stated that text modification did not create a significant difference in general comprehension in B1 level adult Turkish learners. In this study, however, it was observed that simplification performed by paying attention to the theme and aesthetic structure with 7th-grade students within the scope of the native language Turkish course positively affected reading comprehension. This difference suggests that the effectiveness of the intervention is not universal and that the reader profile (L1/L2 status, cognitive development level) and the characteristics of the intervention are critical variables. Therefore, this research presents evidence against Maxwell's findings, showing that the intervention can also be effective in informative texts under certain conditions, and provides important data to this discussion from the context of Turkey.

There are also concerns in the literature that text modification processes may damage the aesthetic texture in the authentic structure of texts and reduce the degree to which the author's style is reflected (Yaşar, 2019; Yücelşen, 2014). Thus, the shortcomings of using modified texts are observed. To mitigate this within the scope of the research, expert opinions were consulted after modification and during text selection. However, the use of authentic texts also provides various benefits. Using authentic literature in conjunction with teaching strategies significantly increases students' motivation, enthusiasm, interest, and passion for reading, while it can also expand their vocabularies and prepare them for college and career readiness (Ciecierski & Bintz, 2015). In the context of foreign language teaching, it is also stated that authentic texts contain more cultural information and can foster cultural awareness to a greater extent (Bérešová, 2015). Furthermore, they can attract interest and motivate by exhibiting the qualities that literary works possess (Illés, 2008). The use of language-specific patterns offered by authentic texts, natural language flow, and aesthetic value are among the most important advantages of these texts.

One of the most central findings of the research is that the effect of text modification is more distinct in informative texts (large effect) compared to narrative texts (medium effect). This finding shows a remarkable parallel with a result reached by Bayraktar and Durukan (2020) in the national literature. Bayraktar and Durukan observed that surface structure changes in informative texts created a significant difference at all student levels (low, medium, high) and determined that the intervention was more functional in the informative text type. This study, on the other hand, also presents a different quantitative comparison by calculating effect sizes. The theoretical reasons underlying this situation can be explained by the difference in cognitive supports that text types offer to the reader. The lower effect size in the narrative text may be related to the fact that narration, by its nature, provides the reader with more opportunities to fill in the gaps regarding the text (Göktürk, 2023). Indeed, Yücelşen's (2014) study supports this idea by showing that shortenings in modified stories do not deviate significantly from their authentic versions in terms of meaning. Since the informative text structure does not allow much for this, it can be said that it is not as favorable as the narrative text in terms of harboring enough clues to compensate for the insufficiency of vocabulary knowledge during reading. This situation may be directly related to the fact that expository texts are more challenging than narrative texts, as stated in the literature (Dickens & Meisinger, 2017). In these texts lacking narrative scaffolds, linguistic barriers may carry the potential to directly transform into a comprehension barrier. Therefore, adaptations targeted at such texts may provide a proportionally greater gain. This situation can be associated with a discussion in the scientific text adaptation literature. For example, Norris et al. (2012) compared two different adaptations of a scientific article: Adapted Primary Literature (APL), which preserves the scientific structure of the authentic article, and Secondary Literature (SL), written in a more narrative style. It was found that SL increased comprehension, while the APL format developed critical thinking

more. The intervention in this study is structurally close to the APL type and shows that such an adaptation creates a strong effect on basic comprehension as well. This situation can be explained by the features of an effective APL defined by Ariely et al. (2019): Preserving the authentic structure of scientific writing while reducing lexical complexity. Therefore, the findings of this study indicate that an adaptation in accordance with APL principles can be an effective approach in terms of comprehension as well.

According to findings, the effect of text modification (simplification, elaboration, explication) on reading comprehension is observed. However, when the resulting score difference is evaluated, it can be stated that text modification should not be a process frequently resorted to at the seventh-grade level. Because although a significant difference occurred, this difference was determined as ~2 points out of maximum scores of 16 and 18. Conducting text modification manually may also not be efficient in terms of time and effort (Uygur, 2024). While manual text modification is viewed as an effective method, developments in this field have brought about automation (Siddharthan, 2014). In this context, the creation of automated text modification systems specific to Turkish may carry the potential to produce texts with high meaning and cohesion. One study suggests that the decline in SAT-verbal scores can be attributed to the cumulative deficit in domain-specific knowledge caused by long-term exposure to simplified school textbooks, which affects students' reading comprehension and verbal achievement levels (Hayes et al., 1996). As can be seen, exposing students to simpler texts for a long time may yield negative results. Additionally, the characteristics of text complexity affect the type of comprehension processes middle school readers form while reading science texts. In the study conducted by Dahl et al. (2021), it was revealed that middle school students made more detailed inferences in science texts containing simple syntax, deep cohesion, and concrete words; whereas they developed more metacognitive and abstract evaluations in complex science texts. In this research, texts were simplified at the level of syntactic structure and vocabulary, but an effort was made to preserve the cohesion structure. It can be said that students exhibiting higher performance in comprehension with texts arranged in this way shows similarity with that study. Furthermore, as can be seen, text complexity affects students' comprehension processes. Here, acting by considering the areas educators need to focus on can be evaluated as important.

Consequently, the findings of this research have revealed that text modification is not a uniform intervention that works to the same extent in every situation, and its effectiveness varies significantly depending on the structural characteristics of the text. This situation suggests that the natural narrative scaffolds (plot, character, etc.) possessed by narrative texts can serve as a kind of cognitive buffer for readers in overcoming linguistic difficulties, but that linguistic accessibility plays a much more critical role in informative texts that have high conceptual density and lack these scaffolds. Thus, it can be stated that in order to understand the success of the intervention, the internal features of the text and the cognitive supports it offers to the reader need to be examined in more depth.

Limitations

There are some methodological limitations that need to be considered when interpreting the findings of this study. The fact that the research has a quasi-experimental design and participants were not randomly assigned may imply that there could be unobserved differences between groups, even though initial reading comprehension levels were controlled with ANCOVA. Additionally, the fact that participants came from different schools might have caused uncontrolled school and teacher effects to play a potential role in the results. Another methodological limitation can be stated as the pre-test measuring general reading comprehension skills (multiple-choice, open-ended, etc.), while the post-test measured only the texts specific to the intervention. In terms of the scope of the research, working with only two different text types, the limited number of students in the study group, and the failure to control for individual reading strategies used by students during applications as a variable restrict the generalizability of the findings. Finally, the fact that the applied text modification intervention included multiple techniques (lexical simplification, syntactic simplification, elaboration, etc.) together prevents determining how effective each technique is on its own.

Recommendations

In line with the findings obtained from this research and the limitations identified, the following recommendations can be offered:

- In this study, multiple modification techniques were used together with a holistic approach. Moving from this limitation, future studies should investigate the effects of only lexical simplification, only syntactic simplification, or elaboration on reading comprehension separately or comparatively to research which technique is more effective in the Turkish context.
- To overcome limitations such as the quasi-experimental design of this study and recruiting participants from different schools, conducting future research with true experimental designs with random assignment and in environments where external variables such as school/teacher are controlled will increase the internal validity of the findings.
- Finally, to go beyond the quantitative effect of text modification on reading comprehension and to understand the role of uncontrollable variables such as reading strategies, methods such as eye-tracking or think-aloud protocols can be used. This will offer deeper answers to the question "why and how does text modification work?"

Implications for Practice

The results obtained from this research allow us to offer concrete and evidence-based implications for the field of education regarding text modification:

- First, it was observed that text modification significantly increased reading comprehension, but this increase corresponded to approximately 12-13% of the total score. The most fundamental implication of this situation is that performing text modification manually may not always be an economical process in terms of time and effort. Therefore, the priority strategy for teachers and curriculum developers should be presenting students directly with linguistically accessible authentic texts suitable for their age and cognitive levels. Text adaptation, on the other hand, should be viewed as a targeted support tool, especially for mandatory historical or foundational texts.
- The most critical finding of the study is that the effect of modification is larger in informative texts (η^2_p = .141) compared to narrative texts (η^2_p = .087). The implication of this finding for textbook authors and curriculum developers could be as follows: The linguistic structure of texts, especially in fields with high abstract and conceptual density such as Science and Social Studies, is at least as important as their content. Subjecting texts in these fields to a rigorous adaptation process in terms of both theme and word frequency and sentence complexity can affect student achievement by increasing the level of comprehension.

References

Aksan, Y., Aksan, M., Mersinli, Ü., & Demirhan, U. U. (2016). *A frequency dictionary of Turkish*. Routledge. https://doi.org/10.4324/9781315733302

Al-Thanyyan, S., & Azmi, A. (2021). Automated text simplification. *ACM Computing Surveys (CSUR), 54*(2), 1-36. https://doi.org/10.1145/3442695

Amendum, S., Conradi, K., & Hiebert, E. (2018). Does text complexity matter in the elementary grades? A research synthesis of text difficulty and elementary students' reading fluency and comprehension. *Educational Psychology Review*, (30), 121-151. https://doi.org/10.1007/s10648-017-9398-2

American Educational Research Association. (2011). *Code of ethics*. https://www.aera.net/About-AERA/AERA-Rules-Policies/Professional-Ethics

American Psychological Association. (2020). *Publication manual of the American Psychological Association* (7th ed.). https://apastyle.apa.org/products/publication-manual-7th-edition

Arfé, B., Mason, L., & Fajardo, I. (2018). Simplifying informational text structure for struggling readers. *Reading and Writing,* (31), 2191-2210. https://doi.org/10.1007/s11145-017-9785-6

Ariely, M., Livnat, Z., & Yarden, A. (2019). Analyzing the language of an adapted primary literature article. *Science & Education*, (28), 63-85. https://doi.org/10.1007/S11191-019-00033-5

Arslan, A., & Engin, A. O. (2019). 5. sınıf Türkçe ders kitabının öğretmen görüşlerine göre değerlendirilmesi. *Uluslararası Türk Eğitim Bilimleri Dergisi, 2019*(13), 67-94. https://doi.org/10.46778/goputeb.570245

Ateşman, E. (1997). Türkçede okunabilirliğin ölçülmesi. *Ankara Üniversitesi TÖMER Dil Dergisi*, (58), 171-174.

Aydın, G., & Ayrancı, B. (2018). Reading preferences of middle school students. *World Journal of Education*, *8*(2), 127-139. https://doi.org/10.5430/WJE.V8N2P127

Aytan, T., Saydam M., Akkaş, N., & Güneş, G. (2021). Yabancılara Türkçe öğretiminde sadeleştirilmiş metinlerin okuduğunu anlamaya etkisi: Nasrettin Hoca Fıkraları örneği. *International Journal of Languages Education and Teaching, 9*(1), 230-242. https://doi.org/10.29228/ijlet.49240

Baddeley, A. (2003). Working memory: Looking back and looking forward. *Nature Reviews Neuroscience*, *4*(10), 829-839. https://doi.org/10.1038/nrn1201

Bağcı-Dağdeviren, Ş. (2024). *İlkokul öğrencileri için düzeyleştirilmiş metin ve okunabilirlik formülü geliştirilmesi üzerine bir araştırma* (Thesis No. 850317) [Doctoral dissertation, İstanbul University]. Council of Higher Education National Thesis Center.

Bakan, H. (2012). *Yabancılara Türkçe öğretiminde metindilbilimsel ölçütler çerçevesinde bir sadeleştirme denemesi: Sait Faik Abasıyanık, Meserret Oteli* (Thesis No. 304267) [Master's thesis, Ankara University]. Council of Higher Education National Thesis Center.

Bakan-Aktaş, H., & Ay, S. (2021). Yabancı dil öğretiminde geleneksel metin sadeleştirme yaklaşımına karşı metindilbilimsel sadeleştirme önerisi: Bağlaşıklık ve bağdaşıklık kavramları. *Ankara Üniversitesi Sosyal Bilimler Dergisi, 12*(1), 94-102. https://izlik.org/JA68ZC96UX

Başaran, M. (2021). Okuduğunu anlayamayan öğrencilerin okuma esnasındaki bilişsel davranışları ve duygu durumları. *Ana Dili Eğitimi Dergisi, 9*(1), 45-58. https://doi.org/10.16916/aded.802475

Bayraktar, İ., & Durukan, E. (2016). Türkçe ders kitaplarında metin sadeleştirme üzerine bir değerlendirme. *Uluslararası Türkçe Edebiyat Kültür Eğitim Dergisi, 5*(3), 1356-1369. https://doi.org/10.7884/teke.709

Bayraktar, İ., & Durukan, E. (2020). Yüzey yapı değişiklikleri ile öğrencilerin okuduğunu anlama becerileri arasındaki ilişki üzerine bir inceleme. *Ana Dili Eğitimi Dergisi*, *8*(2), 292-314. https://doi.org/10.16916/aded.668524

Beck, I. L., McKeown, M. G., Omanson, R. C., & Pople, M. T. (1984). Improving the comprehensibility of stories: The effects of revisions that improve coherence. *Reading Research Quarterly*, *19*(3), 263-277. https://doi.org/10.2307/747821

Bérešová, J. (2015). Authentic materials - enhancing language acquisition and cultural awareness. *Procedia - Social and Behavioral Sciences*, (192), 195-204. https://doi.org/10.1016/J.SBSPRO.2015.06.028

Bölükbaş, F. (2015). Yabancı dil olarak Türkçe öğretiminde okuma metinlerinin dil düzeylerine göre sadeleştirilmesi. *International Journal of Languages' Education and Teaching,* (UDES2015), 924-935.

Brookhart, S. M. (2010). *How to create and use rubrics for formative assessment and grading*. ASCD.

Carlisle, J. (2000). Awareness of the structure and meaning of morphologically complex words: Impact on reading. *Reading and Writing*, (12), 169-190. https://doi.org/10.1023/A:1008131926604

Carroli, P. (2008). *Literature in second language education: Enhancing the role of texts in learning*. Bloomsbury Academic. https://doi.org/10.5040/9781474212267

Cerga-Pashoja, A., Gaete, J., Shishkova, A., & Jordanova, V. (2019). Improving reading in adolescents and adults with high-functioning autism through an assistive technology tool: A cross-over multinational study. *Frontiers in Psychiatry*, (10), 546. https://doi.org/10.3389/fpsyt.2019.00546

Ciecierski, L., & Bintz, W. (2015). Using authentic literature to develop challenging and integrated curriculum. *Middle School Journal*, *46*(5), 17-25. https://doi.org/10.1080/00940771.2015.11461921

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Lawrence Erlbaum Associates.

Creswell, J. W. (2020). *Eğitim araştırmaları* (H. Ekşi, Ed. & Trans.). Edam.

Çetinkaya, G. (2010). *Türkçe metinlerin okunabilirlik düzeylerinin tanımlanması ve sınıflandırılması* (Thesis No. 265580) [Doctoral dissertation, Ankara University]. Council of Higher Education National Thesis Center.

Dahl, A., Carlson, S., Renken, M., McCarthy, K., & Reynolds, E. (2021). Materials matter: An exploration of text complexity and its effects on middle school readers' comprehension processing. *Language, Speech, and Hearing Services in Schools*, *52*(2), 702-716. https://doi.org/10.1044/2021_LSHSS-20-00117

Daneman, M., & Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior*, *19*(4), 450-466. https://doi.org/10.1016/S0022-5371(80)90312-6

De Belder, J., & Moens, M. F. (2010). Text simplification for children. In F. Crestani, S. Marchand-Maillet, H. Chen, E. N. Efthimiadis, & J. Savoy (Eds.), *Proceedings of the 33rd International ACM SIGIR Conference on Research and Development in Information Retrieval* (pp. 19-26). Association for Computing Machinery.

Derse, G. E., & Coşkun, H. (2021). Türkçe ders kitabındaki metinlerin değerler eğitimi açısından incelenmesi ve öğretmen görüşlerine göre değerlendirilmesi. *Değerler Eğitimi Dergisi*, *19*(41), 9-35. https://doi.org/10.34234/ded.738543

Dickens, R., & Meisinger, E. (2017). Examining the effects of reading modality and passage genre on reading comprehension in middle school students. *Reading Psychology*, *38*(3), 321-347. https://doi.org/10.1080/02702711.2016.1263701

Duman, A. (2010). Türkçe eğitiminde metne müdahale sorunu. *Türklük Bilimi Araştırmaları*, (27), 285-295. https://dergipark.org.tr/en/download/article-file/156994

Durmuş, M. (2013a). İkinci/yabancı dil öğretiminde sadeleştirilmiş metin sorunları. *Bilig*, (65), 135-150. https://bilig.yesevi.edu.tr/yonetim/icerik/makaleler/2422-published.pdf

Durmuş, M. (2013b). Metin değiştiriminin dilbilimsel süreçleri üzerine. *The Journal of Academic Social Science Studies*, *6*(4), 391-408. https://doi.org/10.9761/JASSS1077

Durukan, E. (2014). Metinlerin okunabilirlik düzeyleri ile öğrencilerin okuma becerileri arasındaki ilişki. *Ana Dili Eğitimi Dergisi*, *2*(3), 68-76. https://doi.org/10.16916/aded.26659

Elleman, A., Lindo, E., Morphy, P., & Compton, D. (2009). The impact of vocabulary instruction on passage-level comprehension of school-age children: A meta-analysis. *Journal of Research on Educational Effectiveness*, *2*(1), 1-44. https://doi.org/10.1080/19345740802539200

Field, A. (2024). *Discovering statistics using IBM SPSS Statistics*. SAGE Publications.

Fleiss, J. L. (1971). Measuring nominal scale agreement among many raters. *Psychological Bulletin, 76*(5), 378-382. https://doi.org/10.1037/h0031619

Foorman, B., Petscher, Y., & Bishop, M. (2012). The incremental variance of morphological knowledge to reading comprehension in grades 3-10 beyond prior reading comprehension, spelling, and text reading efficiency. *Learning and Individual Differences*, *22*(6), 792-798. https://doi.org/10.1016/J.LINDIF.2012.07.009

Gala, N., François, T., Javourey-Drevet, L., & Ziegler, J. C. (2018). La simplification de textes, une aide à l'apprentissage de la lecture. *Langue Française, 199*(3), 123-131. https://doi.org/10.3917/lf.199.0123

Gamer, M., Jim Lemon, J., & Singh, I. (2019). *irr: Various coefficients of interrater reliability and agreement* (Version 0.84.1) [R package]. https://CRAN.R-project.org/package=irr

Gilmore, A. (2007). Authentic materials and authenticity in foreign language learning. *Language Teaching, 40*(2), 97-118. https://doi.org/10.1017/S0261444807004144

Glavas, G., & Štajner, S. (2015). *Simplifying lexical simplification: Do we need simplified corpora?*. In C. Zong & M. Strube (Eds.), *Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing (Volume 2: Short Papers)* (pp. 63-68). Association for Computational Linguistics. https://doi.org/10.3115/v1/P15-2011

Göktürk, A. (2023). *Okuma uğraşı.* Yapı Kredi Yayınları.

Güneş, F. (2013). Türkçe öğretiminde metin seçimi. *Ana Dili Eğitimi Dergisi, 1*(1), 1-12. https://doi.org/10.16916/aded.16014

Hayes, D. P., Wolfer, L. T., & Wolfe, M. F. (1996). Schoolbook simplification and its relation to the decline in SAT-verbal scores. *American Educational Research Journal*, *33*(2), 489-508. https://doi.org/10.3102/00028312033002489

Illés, É. (2008). What makes a coursebook series stand the test of time. *Elt Journal*, *63*(2), 145-153. https://doi.org/10.1093/ELT/CCN019

İsen, D., & Işınsu, T (2011). Klasik eserlere çocuklar nasıl yönlendirilmeli ve buna ilişkin batıdan örnekler. *Tuskish Studies, 6*(3), 929-937. http://doi.org/10.7827/TurkishStudies.2552

Javourey-Drevet, L., Dufau, S., François, T., Gala, N., Ginestié, J., & Ziegler, J. (2022). Simplification of literary and scientific texts to improve reading fluency and comprehension in beginning readers of French. *Applied Psycholinguistics*, *43*(2), 485-512. https://doi.org/10.1017/S014271642100062X

Kanık-Uysal, P. (2022). Türkçe dersi sınavlarında yer alan soruların üst düzey düşünme becerileri açısından incelenmesi. *Ana Dili Eğitimi Dergisi, 10*(1), 136-156. https://doi.org/10.16916/aded.1016659

Karaağaç, H. G., & Alikılıç, D. (2023). Çocuk edebiyatı ve çocuğa görelik kavramı. *Education and Technology in Information Science*, *1*(1), 58-66. https://izlik.org/JA22TP34TR

Kim, Y. S., & Snow, C. E. (2009). Text modification: Enhancing English language learners' reading comprehension. In E. H. Hiebert & M. Sailors (Eds.), *Finding the right texts: What works for beginning and struggling readers* (pp. 129-146). Guilford Press.

Kintsch, W. (1998). *Comprehension: A paradigm for cognition*. Cambridge University Press.

Krippendorf, K. (2011). *Computing Krippendorff Alpha-Reliability*. https://www.asc.upenn.edu/sites/default/files/2021-03/Computing%20Krippendorff%27s%20Alpha-Reliability.pdf

Kurudayıoğlu, M., & Soysal, T. (2019). 2018 Türkçe Dersi Öğretim Programı kazanımlarının 21. yüzyıl becerileri açısından incelenmesi. *Ahi Evran Üniversitesi Sosyal Bilimler Enstitüsü Dergisi*, *5*(2), 483-496. https://doi.org/10.31592/aeusbed.621132

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, *33*(1), 159-174. https://doi.org/10.2307/2529310

Li, Y., Xu, S., & Wang, Q. (2005). The effects of simplified and elaborated texts on second language reading comprehension: An exploratory study. *Vigo International Journal of Applied Linguistics*, (2), 45-74.

Long, M. H., & Ross, S. (1993). Modifications that preserve language and content. In Tickoo, M. L. (Ed.), *Simplification: Theory and application* (Anthology Series No. 31, pp. 29-52). SEAMEO Regional Language Centre.

Maxwell, S. (2011). *The effects of two types of text modification on English language learners reading comprehension: simplification versus elaboration* (Publication No. 449) [Master's thesis, Hamline University]. School of Education and Leadership Student Capstone Theses and Dissertations.

Melanlıoğlu, D. (2008). Kültür aktarımı açısından Türkçe öğretim programları. *Eğitim ve Bilim Dergisi, 33*(150), 64-73. https://educationandscience.ted.org.tr/article/view/785

Mesmer, H., & Hiebert, E. (2015). Third graders' reading proficiency reading texts varying in complexity and length. *Journal of Literacy Research*, *47*(4), 473-504. https://doi.org/10.1177/1086296X16631923

Ministry of National Education. (2006). *İlköğretim (6, 7, ve 8. sınıflar) Türkçe dersi öğretim programı.* Talim ve Terbiye Kurulu Başkanlığı.

Ministry of National Education. (2019). *Türkçe dersi öğretim programı.* Talim ve Terbiye Kurulu Başkanlığı.

Norris, S., Stelnicki, N., & Vries, G. (2012). Teaching mathematical biology in high school using adapted primary literature. *Research in Science Education*, (42), 633-649. https://doi.org/10.1007/s11165-011-9215-8

Özcan, H. (2008). Halk edebiyatı metinlerinin çocuk edebiyatına kaynak olması ve örneklem olarak Dede Korkut Hikayeleri. *Turkish Studies*, *3*(2), 582-603. http://doi.org/10.7827/TurkishStudies.311

Özçelik, D. A. (2016). *Okulda ölçme ve değerlendirme* (5th ed.). Pegem Akademi.

Reed, D., & Kershaw-Herrera, S. (2016). An examination of text complexity as characterized by readability and cohesion. *The Journal of Experimental Education*, *84*(1), 75-97. https://doi.org/10.1080/00220973.2014.963214

Seol, H. (2024). *Seolmatrix: Correlations suite for jamovi*. (Version 3.8.8) [Computer software]. https://github.com/hyunsooseol/seolmatrix

Sever, S. (2018). *Sanatsal uyaranlarla dil öğretimi*. Tudem.

Siddharthan, A. (2014). A survey of research on text simplification. *ITL-International Journal of Applied Linguistics, 165*(2), 259-298. https://doi.org/10.1075/itl.165.2.06sid

Spencer, M., Gilmour, A., Miller, A., Emerson, A., Saha, N., & Cutting, L. (2018). Understanding the influence of text complexity and question type on reading outcomes. *Reading and Writing*, (32), 603-637. https://doi.org/10.1007/s11145-018-9883-0

Surya, S., Mishra, A., Laha, A., Jain, P., & Sankaranarayanan, K. (2018). *Unsupervised neural text simplification* [Preprint]. arXiv. https://arxiv.org/abs/1810.07931

Sweller, J., Van Merriënboer, J. J., & Paas, F. (2019). Cognitive architecture and instructional design: 20 years later. *Educational Psychology Review*, (31), 261-292. https://doi.org/10.1007/s10648-019-09465-5

Şeref, İ., Karagöz, B., & Koç Ardıç, İ. (2020). Türkçe ders kitaplarındaki okuma metinlerinin metinsellik özellikleri: Lao Tzu'nun 'Acele Karar Vermeyin' öyküsü örneği. *Elektronik Sosyal Bilimler Dergisi, 19*(73), 92-104. https://doi.org/10.17755/esosder.516826

Şimşek, T. (2007). *Çocuk edebiyatı*. Suna Yayınları.

Tekşan, K., & Çinpolat, E. (2023). Türkçe ders kitaplarındaki değiştirilmiş anlatı metinlerine ilişkin bir inceleme. *Erzincan Üniversitesi Eğitim Fakültesi Dergisi, 25*(3), 416-427. https://doi.org/10.17556/erziefd.1245443

Tekşan, K., Çinpolat, E., & Uyar, Y. E. (2023). 2000-2022 yılları arasında kullanılan Türkçe ders kitaplarında metin seçimleri. *Gazi Eğitim Fakültesi Dergisi, 43*(3), 2129-2178. https://doi.org/10.17152/gefad.1246579

Tezcan-Aksu, B., & Adalı, E. (2018). *Çağdaş Türkçenin sıklık sözlüğü*. Ötüken Neşriyat.

The Jamovi Project. (2023). *Jamovi* (Version 2.4) [Computer software]. https://www.jamovi.org

Uygur, M. E. (2024). Dilbilimsel bakış açısıyla uygulamalı metin değiştirimi. In H. H. Mutlu & E. Çinpolat (Eds.), *Türkçenin yabancı dil olarak öğretiminde metin değiştirimi* (pp. 19-59). Vizetek.

Uygur, M. E., Mutlu, H. H., & Çinpolat, E. (2025). The effect of text modification on comprehension and the exploration of perceptions. *Reading in a Foreign Language, 37*(1), 1-34. https://doi.org/10.64152/10125/67519

Ülper, H., Çetinkaya, G., & Bayat, N. (2017). Okuduğunu anlama testinin geliştirilmesi. *Ahi Evran Üniversitesi Kırşehir Eğitim Fakültesi Dergisi*, *18*(1), 175-190.

Yano, Y., Long, M. H., & Ross, S. (1994). The effects of simplified and elaborated texts on foreign language reading comprehension. *Language Learning*, *44*(2), 189-219. https://doi.org/10.1111/j.1467-1770.1994.tb01100.x

Yaşar, S. A. (2019). *Yabancılara Türkçe öğretiminde metin uyarlama ve Memduh Şevket Esendal'ın Pazarlık hikâyesinin A2 seviyesine göre uyarlanması* (Thesis No. 606467) [Master's thesis, Dumlupınar University]. Council of Higher Education National Thesis Center.

Young, D. J. (1999). Linguistic simplification of SL reading material: Effective instructional practice?. *The Modern Language Journal*, *83*(3), 350-366. https://doi.org/10.1111/00267902.00027

Yücelşen, N. (2014). *Ortaokul Türkçe ders kitaplarında yer alan düzenlenmiş ve kısaltılmış yazınsal metinlerin değerlendirilmesi* (Thesis No. 381937) [Master's thesis, İstanbul University]. Council of Higher Education National Thesis Center.

Appendices

1. Graded Scoring Key for the Narrative Text Scent of Hyacinth (Sümbül Kokusu) Comprehension Questions

Question No	0 Points	1 Point	2 Points	3 Points
1. According to the text, what is the reason for Hüseyin Arif being in Budapest?	No Response / Other	Education/Learning/Reading (Answers stating only education) *Ex: "To get education."*	Completing university education/learning (Answers where the level of education is also given) *Ex: "He is in Budapest to receive education at the university."*	–
2. Give two examples from the text regarding the attitudes of Hüseyin Arif or Mehmet Siyavuş showing their patriotism.	No Response / Other	Sadness upon reading the newspaper / Going to the Gallipoli war / Smelling the hyacinth and remembering the homeland (One attitude given without explanation) *Ex: "Wanting to fight for Çanakkale."*	Sadness upon reading the newspaper and going to the Gallipoli war, etc. (Multiple attitudes given with explanation) *Ex: "Worrying about the homeland and selling everything they have."*	–
3. In which aspect are the British-French and the Turks on the battlefield compared in the text?	No Response / Other	In terms of military technology (Answers where details are not fully given) *Ex: "In terms of military."*	Armored ships, airplanes, and submarines are compared with our chest and our arm (Answers where the compared elements are presented) *Ex: "Compares armored ships, airplanes, and submarines."*	–
4. About which subject is Arif hopeless in the first part of the text?	No Response / Other	About the war (Only the subject is given) *Ex: "About the defense of Istanbul."*	He is hopeless about the course of the war in his homeland because the enemy is very strong (Subject given together with reasons) *Ex: "About how to prevent the attacks coming from land and sea."*	–

5. What is the smelling of the hyacinth flower by the characters a symbol of?	No Response / Other	Homeland/Longing (Presentation of the source of the symbol or only the emotion) *Ex: "Symbol of the smell of Istanbul."*	Symbol of the longing felt for the homeland (Answering by specifying the emotion) *Ex: "It is the symbol of the homeland because it is the smell of their country."*	–
6. What are the dominant emotions in the text?	No Response / Other	Longing / Patriotism / Hopelessness / Unhappiness (Presentation of one emotion/patriotism is a must, others are possible) *Ex: "Patriotism."*	Courage, Longing, Patriotism, Hopelessness, Unhappiness (Presentation of two or more emotions) *Ex: "Longing, courage, patriotism."*	–
7. What is the most comprehensive judgment that can be deduced from the text?	No Response / Other	Answers related to patriotism (Non-judgmental sentences related to the text) *Ex: "Patriotism."*	Answers related to patriotism and longing (Sentences reflecting a part of the text) *Ex: "Patriotism is important."*	Patriotism and longing cannot be hindered (Judgmental sentences reflecting the whole text) *Ex: "We must fight for our country at all costs."*
8. Do you find it right for the characters to leave the country they are studying in and return to the homeland to fight? Why?	No Response / Other	I find it right / I do not find it right (Ideas presented without reason) *Ex: "I find it right."*	I find it right because of the importance of the homeland / I do not find it right, they cannot change anything (Idea based on a single reason) *Ex: "Yes, it is a nice thing that they try to protect their homeland."*	I find it right because patriotism is superior to everything / I do not find it right because their duty is to receive education, they can contribute to the homeland that way, and two people cannot change the war (Answers given based on reasons) *Ex: "I find it right because their own hometown, the houses they were born and raised in, their neighborhoods, and their families are there."*

2. Graded Scoring Key for the Informative Text What is Language (Dil Nedir) Comprehension Questions

Question No	0 Points	1 Point	2 Points	3 Points
1. According to the text, why must humans behave in accordance with the nature of language?	No Response / Other	Because language is a natural tool. *E.g.: "Because language is natural."*	Because language is not an artificial or material tool, it is a natural tool. *Ex: "Because language is a natural tool, it is not artificial or material."*	–
2. What could be the purpose of the analogy made with the horse-automobile comparison in the text?	No Response / Other	To emphasize that language is a natural tool. *Ex: "Language is natural, not like a machine."*	To emphasize that one must behave in accordance with the nature of language, unlike artificial tools. *Ex: "We should not treat language like a machine, it is natural."*	–
3. Which features are emphasized in the text by stating that language is a living entity?	No Response / Other	Undergoing change over time / Not being suitable for every external intervention. *Ex: "Language changes over time."*	It is emphasized that changes occur within its structure over time and that individuals and society cannot interfere with this as they wish. *Ex: "Language changes, but people cannot direct this directly."*	–
4. In your opinion, what could be the reasons for different nations expressing the same concepts with words different from other nations?	No Response / Other	Lifestyles / geographies / livelihoods / genetic features / communities they interact with, etc. (Single reason). *Ex: "Because the places they live are different."*	Answers that take two or more of these factors together. *Ex: "Both geographical conditions and cultural structure are different."*	–
5. According to the text, how should interventions on language be, and when can they be made?	No Response / Other	It should be done within the framework of the rules of the language. / It should be done when there is an effect preventing the natural development of the language. *Ex: "Intervention is made if the language deteriorates."*	It should be done within the framework of the rules of the language when there is an effect preventing the natural development of the language. *Ex: "Intervention is made if the language deteriorates, but in accordance with its rules."*	–

6. Explain the expression "Language is a system of secret agreements" based on the text.	No Response / Other	It is not clear when the common vocabulary of nations emerged. *Ex: "It is not clear when it was formed."*	Nations speak of different concepts with different words. *Ex: "Every nation uses different words."*	Every nation has vocabulary and syntax principles that they use in common among themselves as if they had agreed to meet concepts. It is not known when and how these emerged. *Ex: "It is like a secret agreement as if every nation chose the words among themselves."*
7. Do you agree with the idea that language has a protective effect for a nation? Why? Explain with your reasons.	No Response / Other	As our language becomes alienated with foreign words, our nation also becomes alienated from its essence. / As the features of our language are lost, the values that make us who we are also get lost. *Ex: "It would be bad if our language deteriorates."*	When our language is used more properly, interpersonal relationships become better, and this makes society healthier and stronger. / The more language develops, the more the bond of love and interpersonal relationships develop, and this increases social peace and public security. *Ex: "Yes, if we use our language correctly, society becomes better."*	Since their language is corrupted, people who have difficulty understanding each other constantly misunderstand each other, and this causes various chaos among the people and weakens the nation. / A nation that forgets its own language and starts using other languages may forget who it is after a while and fall under the sovereignty of other nations. *Ex: "If language deteriorates, people won't understand each other, we speak other kinds of things, we won't be united."*

Education and Science

Original Article

Vol 51 (2026) No 226 359-378

The effect of philosophy for children practices on the approaches of children towards aesthetic, art, and music concepts

Peyruze Rana Çetinkaya [1], Tuğçe Ebesek Büyükuğurlu [2]

Abstract

The Philosophy for Children Program (P4C) is an approach based on philosophical inquiry and discussion. This program aims to help children generate ideas within a democratic "community of inquiry." Reflecting on topics of art and aesthetics constitutes an important component of the P4C program. This study aimed to examine the effect of P4C sessions on children's approach towards aesthetic, art, and music concepts. Designed as a single-group pretest-posttest model, this study was conducted with 15 volunteer children attending secondary education across various provinces in Turkey. Data gathered through Word Association Test (WAT), P4C sessions, and interviews with participants were analyzed using thematic analysis, a qualitative data analysis method. Findings indicated that P4C practices increased awareness regarding art and aesthetics, enriched and deepened expressions, significantly transformed conceptual misconceptions, and had a positive, informative impact on children's artistic ideas.

Keywords

Philosophy for/with children
Aesthetics
Aesthetic inquiry
Art
Music

Article Info

Received: 10.14.2024
Accepted: 06.03.2025
Published Online: 04.30.2026

DOI: 10.15390/ES.2026.2471

Introduction

Since the beginning of human history, human beings, in alignment with their fundamental nature of thinking, have asked questions regarding their identity, life, nature, the world, and beyond. Philosophy, which prioritizes asking the right questions rather than finding definitive answers, has been an important companion on humanity's journey of self-understanding from ancient times to the present day.

The inclination toward philosophical questions stems from the innate curiosity and desire for understanding. Human life is characterized by exploration, particularly in early childhood; everything that exists is new and awaits discovery. Questions are essential tools for these discoveries, and children's questions, as Jaspers (2003) noted, indicate humans' existential orientation toward philosophy. Children's questions often contain profound philosophical thought.

Philosophy for/with Children

Critical thinking, considered an essential characteristic of modern individuals, fundamentally aligns with philosophical thinking. Matthew Lipman, the pioneer of the "Philosophy for Children" approach, argued that this ability should be developed and encouraged during childhood. Lipman founded the Institute for the Advancement of Philosophy for Children and developed a specially

[1] Dokuz Eylül University, Buca Faculty of Education, Department of Fine Arts Education, İzmir, Türkiye, peyruze.rana@gmail.com
[2] Bağımsız Araştırmacı, Türkiye, tugceebesek@gmail.com

designed curriculum to nurture children's natural curiosity, inquiry tendencies, and philosophical thinking abilities (Haynes & Murris, 2011). The Philosophy for Children (P4C) approach has also been referred to as 'Philosophy with Children' (PwC) by various philosophers who have specialized in this field following Lipman.

Socratic inquiry is at the core of P4C—a program that can be considered both a thinking education and a collective thinking process. Known also as the Socratic method, Socratic dialogue, or Socrates' technique, this approach involves analytic thinking and discussions that seek truth through unbiased questions and answers (Rondhuis, 2005). P4C supports respectful engagement with diverse perspectives and assumptions, encouraging peaceful resolution of conflicts arising from differences in opinion. This approach is valuable in fostering a participatory and democratic "Community of Inquiry" (CoE) (McLeod et al., 2020). Such a community enables young children to think critically, creatively, and collaboratively within a supportive environment, express and develop their ideas respectfully in diverse perspectives, thus contributing positively to their identity development (Clark, 2017, cited in McLeod et al., 2020).

The P4C program allows children to develop ideas on all philosophical issues related to the world and life, approach them inquisitively, and integrate these concepts into their own experiences (Oral, 2013). Aesthetics, a field of philosophical inquiry, is a significant theme in P4C practices, explicitly designed to support philosophical discussion and reflection on values that inspire philosophical inquiry. The program supports aesthetic education through multiple texts and diverse methods (Hamrick, 1989). For instance, in the P4C curriculum developed by Matthew Lipman and his colleague Ann Margaret Sharp, aesthetic inquiry is exemplified in the novel "Suki" (Lipman, 1978), where characters reading poetry, visiting art museums, and discussing music provide grounds for aesthetic inquiry (Gregory, 2017).

Aesthetics and Aesthetic Inquiry

Aesthetics, which investigates and questions the nature of artistic beauty, explores the elements and laws of beauty, asking why we find something beautiful (Erinç, 2011). It can be defined as an area of thought and theory examining how individuals respond to artistic beauty. According to Kagan (1982), aesthetics does not solely focus on beauty but also includes values that transform into or derive from beauty, such as gracefulness, sublimity, dramatic, as well as their opposites, like ugliness, banality, or comedy. Hence, aesthetics investigates all these artistic values. Paying attention to different qualities of our experiences—such as beauty, ugliness, elegance, and garishness—is part of the domain of aesthetics (Lone, 2012). Human beings have a fundamental need to experience the meaning and value of life through aesthetic experience (Alexander, 1994). Art, as a source of aesthetic thought, directly addresses humanity's effort to create meaning and render life meaningful.

In P4C practices, which involve discussions about aesthetics, art is employed more to ask questions rather than merely convey truths. Engaging in dialogue and encountering others' thoughts and reactions towards artistic works and aesthetic stimuli are significant for individuals, allowing them to question their own experiences and inclinations and contributing to developing their understanding of art (Santi, 2007). A community of inquiry can help enhance the aesthetic against the non-aesthetic (Leckey, 2017). Learning to recognize and appreciate aesthetics enriches the capacities of both our mind and senses and can elevate us to the peak of humanity (Madeja & Onuska, 1977).

From ancient times, human beings have established aesthetic relationships with their surroundings within civilization and left artistic traces upon the world (Nutbrown, 2013). "Aesthetic experience always exceeds aesthetics [...] it is a manifestation, record, and celebration of a civilization's life, a tool for promoting its development, and simultaneously, the ultimate judgment on its quality." (Dewey, 2005). Providing aesthetic experiences and promoting curiosity may help children see existing beauty and the possibilities of beauty. Children who engage with beauty might be inspired by it and

motivated to create even more beauty (Wilson, 2010). According to Lipman and Dewey, an aesthetic experience includes capabilities such as judgment, interpretation, creative expression, and empathy (Leckey, 2017). P4C practices, providing an aesthetic experience through aesthetic stimuli and fostering an inquiry approach based on curiosity, can support these capabilities.

Questions concerning aesthetics, pervasive throughout all aspects of life, encompass more than simply reflecting upon art forms; they include perceptions of our life experiences and how we feel about them. "Reflecting on questions such as the nature of beauty and ugliness and the relationship between our aesthetic experiences and our emotions can help us better recognize and understand the way things feel to us in everyday life." (Lone, 2012). The significance of aesthetics in education lies in supporting children's artistic potential, curiosity, creativity, thought processes, awareness, and ultimately fostering enriched personalities (Chou et al., 2014). Aesthetic experiences during childhood indirectly become part of aesthetic experiences in adulthood, as all adults were once children; thus, excluding childhood from overall aesthetic experience would be inappropriate (Leddy, 2002).

The idea that humans should possess an aesthetic understanding in general, not only in artistic terms, constitutes the core of aesthetic theory. Lone (2012), emphasizing aesthetic sensitivity, argues that individuals with such sensitivity perceive qualities of art forms in ways others cannot. Similarly, Read (1931) highlights the importance of sensory education, noting that our behavior is influenced by aesthetic experiences, which are inherently sensory. Engaging critically in aesthetic reflection through carefully chosen stimuli as part of a community of inquiry and sustaining this thinking environment can effectively foster aesthetic sensitivity and education.

Aesthetics in Philosophy for Children Practices

Notable examples stand out among the practices of Philosophy for Children, which focus on aesthetic thought. In her study, Liptai (2005) emphasizes the differences created by using not only texts but also art and craft objects, as well as everyday natural objects, as stimuli to initiate aesthetic inquiry. Santi (2007), in a qualitative analysis-based study where six famous artworks were used as stimuli to initiate aesthetic discussions, found results supporting the hypothesis that P4C practices in aesthetics could encourage cognitive and cultural advancement through increased aesthetic awareness. Leckey (2017), using a renowned painting as a stimulus to engage high school students in philosophical discussion, detailed the discussion contents and analyzed these dialogues by coherently comparing them with Dewey's and Lipman's ideas on art and aesthetics.

Research on P4C in Turkey

In Turkey, the importance of P4C studies has increased, supported by postgraduate theses and scientific research. Particularly notable practical studies examine the effects and contributions of P4C practices on children's conceptual achievements; critical, inquisitive, creative thinking, and problem-solving skills (Işıklar & Abalı Öztürk, 2022; Karadağ & Demirtaş, 2018; Pala, 2022; Türksoy, 2020); levels of creativity (Pekkarakaş, 2020; Taş, 2017); attitudes toward human values (Mehdiyev & Tozduman Yaralı, 2020); assertiveness, cooperation, and self-control skills (Okur, 2008); philosophical attitudes, behaviors, and sensitivities (Dirican & Deniz, 2019; Kaya, 2020); thinking and communication skills in life studies classes (Boyraz, 2019); ideas regarding various philosophical themes (Karasu, 2018); intellectual, emotional, and social levels (Akkocaoğlu Çayır & Akkoyunlu, 2016); and levels of philosophical inquiry and answering philosophical questions (Demirtaş et al., 2018).

These studies predominantly focus on thinking skills and philosophical attitudes. However, research specifically focusing on aesthetic thinking, or the potential effects of P4C approaches on thinking and perceptions related to aesthetic, art, and music concepts, was not found. Considering that several studies conducted in Turkey identified noticeable conceptual misconceptions regarding aesthetics across different educational levels (Çetinkaya & Kalay Meydan, 2023; Demirel, 2018; Kılcan & Akbaba, 2014; Özalp, 2020), there is a need for scientific findings from studies designed to raise

awareness about aesthetics and art within a community of inquiry. Since aesthetics is an abstract and complex learning domain, P4C practices may offer a suitable option to address this need with a program compatible with children's playful and joyful nature (Battin, 1994), rather than from a theoretical and didactic perspective. Similarly, these practices could help support the idea that questioning music, as the most abstract form of art, can help people understand it more deeply, see it as an aesthetic experience, and engage with music and art more consciously and meaningfully.

In line with this, this research aims to examine the effects of P4C practices on children's approaches toward aesthetic, art, and music concepts and present them within a thematic framework. This objective emphasizes the idea of systematically integrating art-focused philosophical discussions into the learning environment. It is believed that this study holds practical significance as it provides awareness and examples for art educators, music educators, and classroom teachers who conduct art classes. Additionally, addressing philosophical discussions on abstract artistic concepts scientifically within younger age groups indicates theoretical importance. Based on this, the research question is as follows: What is the effect of P4C practices on children's approaches toward aesthetic, art, and music concepts?

Sub-questions derived from the research problem are as follows:

1. What are the pretest and posttest results of Philosophy for Children practices?
2. What are the themes regarding Philosophy for Children discussions?
3. What are children's views regarding Philosophy for Children practices?

Method

Research Design and Model

This study used a single-group pretest-posttest quasi-experimental design (Creswell, 2013) to investigate the impact of P4C practices on aesthetic thinking processes within a single group. Accordingly, a small participant group was selected. The group size was limited to 15 children, enabling deeper interaction and intellectual progression in philosophical discussions. The research design is presented in Table 1.

Table 1. Research Design

Group	Pretest	Implication	Posttest
G_1	O_1	X	O_2

The application process consists of P4C activities spread over three sessions. A music educator and a P4C trainer planned and conducted these sessions. Research data were collected using the Word Association Test (WAT) as a pretest-posttest tool and through interviews conducted after the posttest.

Limitations of the Study

The study was conducted experimentally, with a small participant group of 15 elementary school students. Therefore, the study has limited generalizability to other individuals within the same age group.

Study Group

The study group consists of 15 children aged 10-12. Among the students, there are 10 girls and 5 boys. The sampling method followed in this research is random purposive sampling. In this method, the researcher specifies characteristics that align with the research purpose and seeks individuals who possess those characteristics. Once a target group is identified, individuals who meet the pre-determined criteria are expected to participate in the research (Johnson & Christensen, 2020). Decisions must be made regarding whom or what to sample and how many individuals to sample (Creswell, 2012). For this study, considering the abstract and complex nature of the concepts to be discussed, the criteria necessitated that participants should be 10-12 years of age, the number of individuals in each group should be limited to 15 owing to the nature of the P4C sessions, and the group should consist of individuals who could regularly attend sessions on specified days and times. Following necessary permissions, the first 15 volunteers who met these criteria formed the sample group. The participants' names were changed to pseudonyms. Additional information about them is presented in Table 2.

Table 2. Demographic Information of Participants

Participants	Gender	Age	School	City
Nehir	Female	10	State School	Tekirdag
Sahra	Female	11	State School	Eskisehir
Cesur	Male	10	State Conservatory	Istanbul
Rüzgar	Male	10	State School	Tekirdag
Leman	Female	10	State Conservatory	Istanbul
Bulut	Male	10	State Conservatory	Istanbul
Yaprak	Female	10	State Conservatory	Istanbul
İlkim	Female	12	State School	Istanbul
Ali	Male	11	State School	Eskisehir
Arya	Female	10	State School	Tekirdag
Eylül	Female	10	State Conservatory	Istanbul
Güneş	Female	12	State School	Hatay
Doğa	Male	10	Private School	Istanbul
Parla	Female	10	State Conservatory	Istanbul
Aslı	Female	11	State School	Eskisehir

Data Collection Tools

Word Association Test (WAT):

To examine the impact of philosophy sessions with children on the study group's perceptions of concepts such as aesthetics, art, and music, a WAT was administered as a pretest and posttest tool for each session.

The WAT, which consists of a list of words from which associations are to be gathered, is also known as a relational experiment. Respondents must associate each word in the list with the first word that comes to mind (Kostova & Radoynovska, 2010). This alternative assessment technique helps reveal individuals' cognitive structures regarding a specific topic or concept. The WAT can be used before and after the application to observe changes in perceptions regarding related concepts; after a period of conceptual development, the test can be repeated to examine the changes and learning that have occurred (Bahar et al., 1999; Balbağ & Kaya, 2019).

In this study, the WAT was prepared separately for the three focus concepts (aesthetics, art, and music) to reveal conceptual connections. First, the relevant concept was written five times in a column, and participants were asked to write five words that the key concept evoked in them within 30 seconds. Then they were expected to form a "related sentence" regarding the key concept. As a holistic and fluent form of expression, this related sentence allows participants to "validate" the five separate words they provided for the key concept. The process allows for the broad observation of perceptions regarding the

concepts and helps researchers identify any misconceptions. The word association test used in the study is presented in Figure 1.

Which "related" concepts does the concept of "MUSIC" evoke for you? Write down the first 5 concepts that come to mind, which you think are related to music, within 30 seconds. Then explain what this concept means to you in one sentence.MUSIC: MUSIC: MUSIC: MUSIC: MUSIC: Related Sentence: …… .

Figure 1. Word Association Test for the Concept of Music

- ***Interview Questions:***

During the research planning phase, the researchers prepared an interview form consisting of closed- and open-ended questions. Following the final session, individual and private interviews were conducted with the students, and each was asked to respond to the following questions: *"Have your thoughts about aesthetics, art, and music changed compared to before the sessions? If so, how would you describe these changes?", "If you were to explain your discussions in the philosophy sessions to someone else, how would you describe them?", "What was the most impactful or memorable thing for you in the sessions?", "What are your views and suggestions regarding the philosophy discussions? How might they be improved?".*

Data Collection Process

1. Word Association Tests, P4C Sessions, and Interviews

The philosophy program with children was planned as three sessions, each conducted online once weekly for 60 minutes. The researcher, a P4C trainer, facilitated the sessions.

In the first session, aesthetics was discussed; in the second, art; and in the third, music. Before each session, the WAT form related to the respective concept was sent to the students as a pretest in an online format, and after each discussion, they were asked to complete the form again as a posttest.

During the first session on aesthetics, three famous works by Kandinsky and Van Gogh and an example of graffiti were shown to the students as stimuli. The discussion was initiated by asking which painting they liked best and why. In the second session on the concept of art, the video "The Painting Elephant" was shown as a stimulus, and the discussion began with a question about whether the images should be considered art. In the final session, where the concept of music was discussed, a song by The Beatles was played as a stimulus, the notes of the song were shown, and a discussion was initiated with the question, "Where is the music?". In the following minutes of the session, John Cage's silent piece 4'33" was presented, and the discussion shaped by this stimulus continued with the listening of a Billie Eilish song structured around the rhythmic pattern of a traffic light. All stimuli used in the sessions and the questions directed to the children are shown in Table 3.

Following the posttest administered after the third session, individual interviews were conducted with each student using the prepared questions to capture their thoughts on the philosophy discussions related to the concepts of aesthetics, art, and music.

Each session was recorded and later transcribed by the researchers.

An example of a discussion focusing on the importance of own will and effort in art from the session discussing the concept of art is as follows:

Facilitator: So, in that case, can only humans create art?

Güneş: If it has been done willingly, even if an elephant does it, it can be art. If it is forced, it is not art.

Nehir: Art cannot be forced. Nevertheless, non-human creatures can also create art.

Rüzgar: Humans are unnecessary; for example, robots can create art. If they want to, they can do it.

Facilitator: Do you think robots have their own will?

Rüzgar: I think they will in the future. If a robot creates art with its own will, it can be considered art. If it is forced or done by command, it is not art.

Nehir: I agree. If we code a robot to do it, it is not art. If a robot gains its own will, it can be art.

Cesur: I think every drawing is art, even if it is by command.

Parla: If something is forced, it cannot be art.

Arya: Robots do not exert effort; they do it rotely, so it is not art.

1.1. Stimuli and Questions Used in Philosophy Sessions with Children

Information regarding the stimuli and questions used in the sessions is presented in Table 3.

Table 3. Implementation Process of P4C Sessions

Session	Concept	Stimuli	Questions
1.	Aesthetics	Images: Vincent Van Gogh/ Starry Night Wassily Kandinsky/ Color Study: Squares with Concentric Circles Wassily Kandinsky/Composition VIII A graffiti	Which of these paintings do you think is the most beautiful? Why do you think that one is the most beautiful? Can everyone find the same things beautiful? What criteria do we use to find something beautiful? If I say a picture or an object is "very aesthetic," what comes to your mind? What do we find aesthetic? Are "aesthetic" and "beautiful" the same or different things?
2.	Art	Video: "The Painting Elephant"	Do you think this is art? Does art contain a distinguishing feature that does not exist in other things? Does a work need to be famous and enduring to be art? Can something be art even if no one likes it? Can a painting that no one has seen be a work of art? What is needed for something to be art? Can only humans create art?

Table 3. Continued

Session	Concept	Stimuli	Questions
3.	Music	Pieces of Music: 1. The Beatles: "All My Loving" 2. John Cage: 4'33'' 3. Billie Eilish: "Bad Guy"	Where do you think music resides? Does pressing any key on an instrument make music? Is there music in sheet music? Is there music in sheet music for someone who cannot read it? Is there music in silence? Is sound necessary for music? Is every sound vibration music? If a composer considers silence as music, can silence be music? Is the idea of whether something is music subjective? Are listeners the criteria for music? What is necessary for something to be considered music? Is rhythm music? Is a rhythmic traffic light sound music? What transforms sounds into music?

Data Analysis

The data obtained from the WAT was analyzed using the content analysis method. According to Krippendorff (2004), content analysis is a research technique that derives valid and repeatable inferences from texts or other meaningful materials based on their contexts. In this analysis method, the words in the research data are segmented into units, and units with similar meanings are coded under the same categories (Krippendorff, 2004; Weber, 1990).

Words provided in response to the key concepts in the WAT were examined, and frequency tables were created to identify how often each word was repeated. Prominent words were visualized using the word cloud technique. This technique, which visually summarizes content analysis results, enables quick understanding and interpretation, even by those not specialized in information visualization (Fronza et al., 2013).

Responses to each key concept in the WAT were categorized and presented in a table with examples. Subsequently, responses given by each student in the pretest and posttest were compared internally to examine expected positive changes in personal contexts.

Dialogues in P4C sessions were analyzed using content analysis, categories were determined based on coding, and examples were presented in a table with sample sentences.

Finally, responses to interview questions were analyzed using content analysis. Repeated expressions were identified and coded, and themes were determined accordingly. The findings were presented in a table, and participant statements were noted.

Results

This section presents research findings on the effects of P4C activities on children's perceptions of aesthetic, art, and musical concepts.

Pretest and Posttest Results of P4C Practices

- ***Findings from the Word Association Test***

The data collected from the pretest and posttest Word Association Tests were examined in two stages: In the first stage, the words associated with each key concept were examined, and frequency

tables were created. Words that appeared three times or more were visualized using word clouds. In the second stage, related sentences were categorized and analyzed.

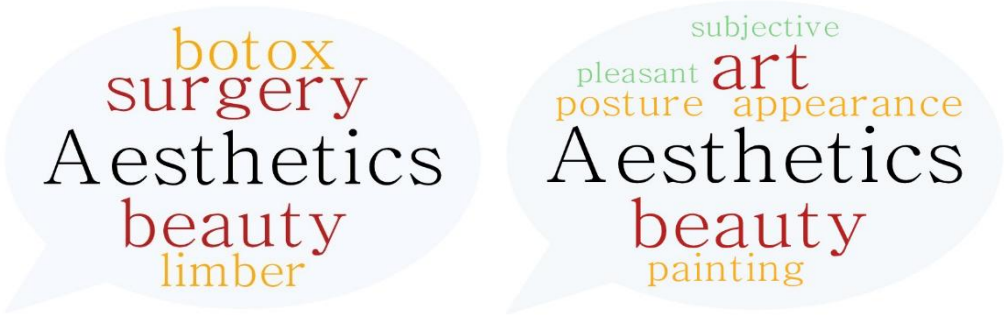


Figure 2. Words Associated with the Concept of Aesthetics in Pretest and Posttest

As seen in Figure 2, the most frequently associated word with "aesthetics" in the pretest was beauty (f=11), followed by surgery (f=6), botox (f=3), and limber (f=3). In the posttest, the concept was associated with beauty (f=13), art (f=6), painting (f=4), posture (f=4), appearance (f=4), subjective (f=3), and pleasant (f=3). In the pretest on aesthetics, the word "surgery," which reflected a misconception and was the most frequently repeated term, was found to be excluded from the network in the posttest. Similarly, words like “botox” and “limber”, which reflected conceptual misunderstandings, were also excluded. Instead, aesthetics became more connected to terms directly related to its philosophical and artistic context, such as “art”, “subjective”, and “pleasant”.

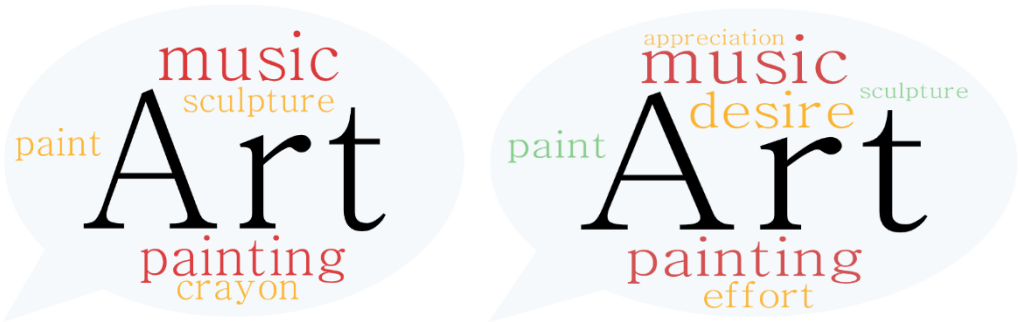


Figure 3. Words Associated with the Concept of Art in Pretest and Posttest

As seen in Figure 3, the words most associated with art in the pretest were painting (f=13), music (f=11), sculpture (f=3), paint (f=3), and crayon (f=3). In the posttest, the associations included painting (f=9), music (f=7), desire (f=4), effort (f=4), appreciation (f=3), paint (f=3), sculpture (f=3). While the pretest words focused mainly on art forms and materials, the posttest associations included terms like “desire” and “effort”, which relate to the creative process, and “appreciation”, which links to aesthetic judgment.

Figure 4. Words Associated with the Concept of Music in Pretest and Posttest

As seen in Figure 4, the most associated words with music in the pretest were note (f=7), instrument (f=7), piano (f=6), song (f=5), art (f=4), sound (f=4), Mozart (f=3), and relax (f=3). In the posttest, the associations shifted to sound (f=12), rhythm (f=7), art (f=7), note (f=5), instrument (f=6), and creativity (f=3). In the pretest, words primarily focused on musical tools and names, whereas in the posttest, emphasis shifted to core musical components like "sound"," rhythm", and "creativity".

- ***Related sentences expressed in the Word Association Test***

The sentences expressed in WAT were examined and classified based on their meanings and informational content. The sentences constructed by participants were categorized into three groups: sentences containing conceptual information, which parallel the given key concept and include a conceptual definition or association; sentences containing personal interpretations or interests related to everyday life and personal views regarding the concept; and sentences containing misconceptions, where expressions differ from or misinterpret the meaning of the key concept within the scope of the research. Frequency tables and sample sentences based on these categories are provided below:

Table 4. Frequency Distribution of Relevant Sentences Categorized by Pretest and Posttest Concerning Key Concepts

Key Concept	Number of sentences containing conceptual information		Number of sentences containing personal opinion/interest		Number of sentences containing misconceptions		Empty responses	
	Pretest	Posttest	Pretest	Posttest	Pretest	Posttest	Pretest	Posttest
Aesthetic	2	9	5	3	7	3	1	-
Art	5	10	10	5	-	-	-	-
Music	5	12	10	3	-	-	-	-

As seen in Table 4, sentences containing conceptual information increase in both pretest and posttest responses for all key concepts, while sentences containing personal opinion/interest decrease. Sentences containing misconceptions are only associated with the aesthetic key concept, and a decrease in such sentences is observed between pretest and posttest responses.

Sample sentences related to the aesthetic key concept are provided in Table 5.

Table 5. Sentences related to the Aesthetic Concept Obtained in Pre-test and Post-test

WAT	Sentences containing conceptual information	Sentences containing personal opinion/interest	Sentences containing misconceptions
Pretest for Aesthetic	"Every art form has a relationship with aesthetics."	"I find ballet very aesthetic." "One's posture while playing the piano is very aesthetic."	"My mother is against getting plastic surgery." "Many people have had cosmetic surgery on their nose." "I think aesthetics is the same as the word 'limber '."
Posttest for Aesthetic	"Aesthetics is a philosophical concept." "Aesthetics is a perspective that can vary from person to person." "Aesthetics can even be liking something that appears dreadful or strange." "The emergence of grace along with aesthetics will add courtesy to society."	"The details in this table are very aesthetic." "This orchestra looks very aesthetic." "The brush strokes in some paintings add a very aesthetic appearance."	"My mother is against getting plastic surgery."

Sample sentences related to the art key concept are provided in Table 6.

Table 6. Sentences Related to the Art Concept Obtained in Pretest and Posttest

WAT	Sentences containing conceptual information	Sentences containing personal opinion/interest	Sentences containing misconceptions
Pretest for Art	"Art is written, visual, and sensory beauty."	"I am interested in the art of painting."	-
	"Music, painting, architecture, and other areas where beautiful works are created and things people love are art."	"Art helped me realize my interests and talents."	
		"I really liked the art he/she created."	
Posttest for Art	"Art is a pursuit that requires effort and hard work."	"I really love the art of painting."	-
	"Art shows itself everywhere in life."		
	"A painting or music piece that has been given effort and diligence is art."		

Sample sentences related to the music key concept are provided in Table 7.

Table 7. Sentences Related to the Music Concept Obtained in Pre-test and Post-test

WAT	Sentences containing conceptual information	Sentences containing personal opinion/interest	Sentences containing misconception
Pretest for Music	"Music consists of organized sounds."	"I really enjoy listening to music."	-
	"Music is an art form."	"I can play the guitar, which is a musical instrument."	
		"I try to keep up with the rhythm of the music."	
Posttest for Music	"Music is an art form that encompasses the composition of sounds and silences."	"The rhythm of this music is very beautiful."	-
	"At the core of music is sound, but not every sound is music."		
	"Rhythm is one of the fundamental elements of music."		
	"Music is an art form heard by the ear and felt by the heart."		

Findings Related to P4C Sessions

Based on the coding of discussions in the P4C sessions, prominent words were identified and categorized, and sample sentences were provided.

Table 8. Coding and Categories Related to the Aesthetic Session

Session 1	Coding	Categories	Sample Sentences
Aesthetics	appreciation impact curiosity subconscious subjective perceive	aesthetic subject	"The reason we find something aesthetic may lie in the curiosity it arouses and its ability to attract our attention." "If we find something aesthetic, it is because we connect it with things embedded in our subconscious." "Only things we can perceive can be considered aesthetic — things we can see or hear."
	colour music painting nature appearance posture artistic abstract	aesthetic object/resource	"Not only things we can see, but also music can be aesthetic if captivating." "Aesthetics is a concept related more to artistic images than people's appearances." "Not only things perceived through the senses but abstract things can also be considered aesthetic."
	elegant pleasant beautiful original harmonious	aesthetic value/qualifiers	"Aesthetic refers to things that appear elegant." "In my opinion, aesthetics has multiple meanings, which is why beauty and aesthetics are both similar and different." "The colors look very harmonious; they impressed me."

According to Table 8, based on the participants' responses and dialogues in the session where the aesthetic concept was discussed, three categories emerged: aesthetic subject, aesthetic object/source, and aesthetic value/qualifications.

Table 9. Coding and Categories Related to the Art Session

Session 2	Coding	Categories	Sample Sentences
Art	appreciation popular beautiful recognition	artwork	"Things that multiple people appreciate can be considered works of art." "This is an artwork because it is beautiful." "If it gains recognition, it is a work of art. What matters is not being liked, but being recognized."
	thought diligence effort desire own will	artist	"Even if no one likes it, the important thing is that the artist can express their thoughts through the artwork." "For something to be art, the artist's desire and effort are necessary." "An act becomes art when carried out through the artist's will. It is not art if it is forced or done by command."

According to Table 9, based on the participants' responses and dialogues during the session where the concept of art was discussed, two categories emerged: artwork and artist.

Table 10. Coding and Categories Related to the Music Session

Session 3	Coding	Categories	Sample Sentences
Music	brain heart love listening understanding ear stick in mind	musical perception	"Music is in the brain." "Music exists in the heart and love of the audience." "When you like a song, it sticks in your mind and you just feel like listening to it and singing it repeatedly."
	note sound rhythm silence melody song composer harmony instrument vocalist	musical terms/concepts	"Music is in the notes." "At the core of music is sound." "Music is not just sound; silence is also part of music." "Not all sounds can be music; a certain rhythm is needed."

According to Table 10, in the session where the concept of music was discussed, two categories emerged based on the participants' responses and dialogues: musical perception and musical terms/concepts.

Findings Related to Views on Philosophy Sessions with Children

The themes derived from participants' responses to the interview questions in the final stage of the research are presented below.

The themes related to the answers to the question 'Have your thoughts about aesthetics, art, and music changed compared to before the sessions? If so, how would you describe these changes?' are presented as a word cloud in Figure 5.

Figure 5. Themes related to the changes observed after the P4C sessions

The following are sample participant responses related to these themes: *"Yes, it did. I only knew aesthetics in the context of surgery. In art, I saw things I had not noticed before. In the music session, I got confused and encountered thoughts I had not considered."* (Nehir), *"I started thinking more philosophically. I used to think more superficially about these concepts. After the philosophy sessions, I started thinking more deeply."* (Cesur), *"With what my friends said and thoughts that came to my mind later, my ideas changed. My way of thinking changed."* (Güneş).

The themes of the responses to the question "If you were to explain your discussions in the philosophy sessions to someone else, how would you describe them?" are presented as a word cloud in Figure 6.

Figure 6. Themes related to the expressions in the P4C sessions

The following are sample participant responses related to these themes: *"I would say that my ideas about these concepts have changed and that I started to think differently."* (İlkim), *"I participated in a philosophy session, and it expanded my thoughts. I learned many things about these concepts and understood the connection between the concept of aesthetics and art."* (Yaprak), *"We had useful and fun discussions about the three concepts."* (Leman).

The themes of the responses to the question "What was the most impactful or memorable thing for you in the sessions?" are presented as a word cloud in Figure 7.

Figure 7. Themes related to the statements about the most memorable aspects of the P4C sessions

The following are sample participant responses related to these themes: *"I wondered whether music is in the ear, brain, or sound. I thought about it and was influenced by this discussion."* (Bulut), *"The idea that music is not only sound but also silence affected me. I used to think aesthetics was just surgery, but I learned that pleasing things are also aesthetics."* (Arya), *"I learned the artistic meaning of aesthetics, that session was beneficial for me. Also, the idea that music can include silence impacted me."* (Leman), *"Aesthetics stayed with me the most. I realized the relationship between aesthetics and beauty."* (Rüzgar).

The themes of the responses to the question "What are your views and suggestions regarding the philosophy discussions?" are presented as a word cloud in Figure 8.

Figure 8. Themes related to the opinions and suggestions regarding P4C practices

The following are sample participant responses related to these themes: *"I think it was an enjoyable and nice event. We could delve even deeper into aesthetics, art, and music. I wish there were more and different concepts and topics covered."* (Parla), *"I wish these sessions were held every day. It was very fun."* (Cesur), *"It was an event I would like to participate in again."* (Doğa).

Discussion and Conclusion

This study aimed to examine philosophical thinking and discussions with children regarding three closely related concepts: aesthetics, art, and music, and how this journey changed their modes of thinking. The findings can be summarized based on word association tests, session observations, and interviews.

Conclusions Based on Pretest and Posttest Results

Before the sessions, aesthetics was most commonly associated with the word "beauty." However, when evaluated in the context of other words and sentences the children produced, it became clear that they were referring to beauty as it is achieved through an aesthetic operation like plastic surgery, which is a widespread conceptual misunderstanding of the concept of aesthetics in their cognitive structures. After the sessions, the children continued associating aesthetics with "beauty," but in a different sense. Their sentences established relationships among aesthetics, art, and philosophy through the concept of what is beautiful, and referred more to artistic beauty rather than that resulting from surgical intervention. While their sentences related to the concept before the sessions largely contained personal interpretations and conceptual misunderstandings, there was a marked increase in sentences displaying conceptual knowledge after the sessions (f=12).

Before the sessions on art and music concepts, children's associations were more object/material-focused. However, after the sessions, it was observed that they began to interpret these concepts from the perspective of their processes of emergence, fundamental elements, and existential meanings. Although a transformation as distinct as that observed in the concept of aesthetics was not evident in the derived words, the sentences constructed before the sessions mostly contained personal interpretations and superficial information. In contrast, after the sessions, sentences involving conceptual knowledge and depth increased in number (f=12).

Conclusions Based on the Themes Emerging from P4C Discussions

In the philosophy session where children discussed the concept of aesthetics, they interpreted aesthetics through concepts related to the aesthetic subject, such as appreciation, perception, and the subconscious; the aesthetic object/source, such as music, painting, nature, and abstraction; and aesthetic values/qualities, such as elegance, beauty, and harmony. In the session where they discussed the concept of art, their interpretations focused on the themes of the artwork and the artist. In the music session, they reflected on how music is perceived by the individual and explored its fundamental elements.

Conclusions Based on Interviews

In one-on-one interviews, most children expressed that they had experienced changes in their thoughts and knowledge about aesthetics, and that their initially superficial thinking had deepened after the sessions. They noted that they encountered different perspectives during the discussions, that their expressions and vocabulary related to these concepts became enriched, and that their curiosity about philosophy arose due to the process. After the sessions, they continued to question and think about the concepts, gained new information, and engaged in respectful discussions with their peers. Moreover, they desired to engage in additional P4C activities and discussions on various concepts and topics.

Based on the results obtained, it can be stated that the "acquisition of new knowledge" expressed by the participants occurred spontaneously during the P4C sessions; due to the nature of the process, the facilitator did not explicitly or implicitly present a definition or perspective regarding any concept. Participants' learning and awareness of a concept developed due to the group's collective discussion and inquiry process. All these inferences were made within a limited framework defined by the selected participant group, the period in which the research was conducted, and the cities/regions in which the participants were located; therefore, the generalizability of the findings is limited. It can be said that the data obtained from the research contributes to the literature. Still, there is a need for further research with larger, diverse samples and different methodologies to support these findings.

In a final statement on aesthetics, Eylül remarked, “The emergence of grace with aesthetics will add courtesy to society.” Philosophy can contribute to society by noting beauty through aesthetics. To understand an aesthetic experience, it is necessary to think about it systematically, i.e., to philosophize. For this, one needs to know how to philosophize and make philosophical practices a habit (Marquez, 2017). In this sense, learning to think and create opportunities and spaces is a valuable approach. The P4C program can facilitate creating this space in children’s worlds, as Güneş mentioned in the interviews, “Despite our young age, it was nice to think about such abstract concepts.”

Parsons (1994) mentions the necessity of integrating aesthetics into art classes by arguing that images given to children as stimuli can support their thinking and discussions about beauty, what we find unattractive, and the goals of art. Similarly, Nutbrown (2013) suggests that the aesthetic need, described as an innate human need, should be considered in educational programs and that children’s artistic experiences need to be nurtured for artistic development.

The aesthetic approach is not exclusive to the arts and artists; it is a mental capacity rooted in human intelligence that can be developed (Clough, 2002, as cited in Nutbrown, 2020). To achieve the idea of an aesthetic society created by adults, integrating aesthetics into the world of children, who will become those aesthetic adults, should be a priority. To enable children to become aesthetically mature individuals who can understand, interpret, evaluate, and critically analyze artworks within a specific context (Parsons, 1994), they need an environment and experiences from which they can acquire these skills.

Suggestions

In future studies on Philosophy for/with Children practices related to aesthetics, art, and music, the following could be suggested:

- Investigating the effectiveness over a more extended period by increasing the number of P4C sessions focused on aesthetic and artistic thinking,
- Examining the effects on different age groups and educational levels,
- Raising awareness among art educators regarding P4C practices to bring opportunities for inquiry and critical thinking about art and music into the classroom, and designing in-service training programs for educators.
- Planning seminars, workshops, and projects related to this approach for prospective teachers studying in the Fine Arts Education and Primary Education departments could be recommended.

Acknowledgements

We want to sincerely thank our “Little Thinkers”, who are the participants in our study.

References

Akkocaoğlu Çayır, N., & Akkoyunlu, B. (2016). Çocuklar için felsefe eğitimi üzerine nitel bir araştırma. *Turkish Online Journal of Qualitative Inquiry, 7*(2), 97-133. https://doi.org/10.17569/tojqi.91449

Alexander, T. (1994). Educating the democratic heart: Pluralism, traditions, and the humanities. *Studies in Philosophy and Education,* (13), 243-259. https://doi.org/10.1007/BF01077682

Bahar, M., Johnstone, A. H., & Sutcliffe, R. G. (1999). Investigation of students' cognitive structure in elementary genetics through word association tests. *Journal of Biological Education, 33*(3), 134-142. http://doi.org/10.1080/00219266.1999.9655653

Balbağ, N. L., & Kaya, M. F. (2019). Fourth grade students' cognitive structures regarding "values": Application of the word association test. *Review of International Geographical Education Online, 9*(1), 193-208. https://doi.org/10.33403/rigeo.511574

Battin, M. P. (1994). Cases for kids: Using puzzles to teach aesthetics to children. *The Journal of Aesthetic Education, 28*(3), 89-104. https://doi.org/10.2307/3333403

Boyraz, C. (2019). *İlkokulda çocuklarla felsefe uygulamaları: Bir eylem araştırması*. (Thesis No. 604901) [Doctoral dissertation, Anadolu University]. Council of Higher Education National Thesis Center.

Chou, M., Huang, P., & Yang, C. (2014). Same theory, different day: Inquiry into preschool children's multiple intelligence and aesthetics ability. *Procedia-Social and Behavioral Sciences*, (143), 534-541. https://doi.org/10.1016/j.sbspro.2014.07.432

Creswell, J. W. (2012). *Qualitative inquiry and research design: Choosing among five approaches*. Sage publications.

Creswell, J. W. (2013). *Research design: Qualitative, quantitative, and mixed methods approaches*. Sage publications.

Çetinkaya, P. R., & Kalay Meydan, S. (2023). Study of cognitive structures of professional music education students regarding the concept of aesthetics. *TAY Journal, 7*(3), 827-845. https://doi.org/10.29329/tayjournal.2023.610.06

Demirel, İ. N. (2018). 9-12 yaş çocukların estetik değer algılarının incelenmesi. *Inonu University Journal of the Faculty of Education, 19(2),* 312-327. https://doi.org/10.17679/inuefd.346720

Demirtaş, V. Y., Karadağ, F., & Gülenc, K. (2018). Okul öncesi dönemdeki çocukların felsefi sorgulama süreçlerinde oluşturdukları soruların düzeyi ve verdikleri cevapların niteliği: Çocuklarla felsefe eğitimi. *International Online Journal of Educational Sciences, 10*(2), 277-294. https://doi.org/10.15345/iojes.2018.02.019

Dewey, J. (2005). *Art as experience*. Penguin Publishing Group.

Dirican, R., & Deniz, U. (2019). Felsefe etkinliklerinin okul öncesi dönem çocukların felsefi tutum ve davranışlarına etkisinin incelenmesi. *Kastamonu Eğitim Dergisi, 28*(1), 421-430. https://doi.org/10.24106/kefdergi.3625

Erinç, M. S. (2011). *Sanat Psikolojisi'ne giriş*. Ütopya Yayınevi.

Fronza, I., Janes, A., Sillitti, A., Succi, G., & Trebeschi, S. (2013). Cooperation wordle using preattentive processing techniques. In *6th International Workshop on Cooperative and Human Aspects of Software Engineering (CHASE)* (pp. 57-64). IEEE. https://doi.org/10.1109/CHASE.2013.6614732

Gregory, M. (2017). Philosophy for children: Where are we now?. In S. Naji & R. Hashim (Eds.), *History, theory and practice of philosophy for children* (pp. 137-144). Routledge Press. https://doi.org/10.4324/9781315208732-24

Hamrick, W. S. (1989). Philosophy for children and aesthetic education. *Journal of Aesthetic Education, 23*(2), 55-67. https://doi.org/10.2307/3332941

Haynes, J., & Murris, K. (2011). The provocation of an epistemological shift in teacher education through philosophy with children. *Journal of Philosophy of Education, 45*(2), 285-303. https://doi.org/10.1111/j.1467-9752.2011.00799.x

Işıklar, S., & Abalı Öztürk, Y. (2022). The effect of philosophy for children (P4C) curriculum on critical thinking through philosophical inquiry and problem solving skills. *International Journal of Contemporary Educational Research, 9*(1), 130-142. https://doi.org/10.33200/ijcer.942575

Jaspers, K. (2003). *Way to wisdom: An introduction to philosophy* (2nd ed.). Yale University Press. https://doi.org/10.12987/9780300163575

Johnson, R., & Christensen, L. (2020). *Educational research: Quantitative, qualitative, and mixed approaches*. (7th ed.) Sage Publications.

Kagan, M. (1982). *Güzellik bilimi olarak estetik ve sanat* (A. Çalışlar, Trans.). Altın Kitaplar Yayınevi.

Karasu, F. Z. (2018). *İlkokul çocuklarıyla felsefe yapmak üzerine nitel bir çalışma: Dördüncü sınıf örneği* (Thesis No. 537665) [Master's thesis, Hacettepe University]. Council of Higher Education National Thesis Center.

Karadağ, F., & Demirtaş, V. Y. (2018). The effectiveness of the philosophy with children curriculum on critical thinking skills of pre-school children. *Education and Science, 43*(195), 19-40. https://doi.org/10.15390/EB.2018.7268

Kaya, B. N. (2020). *Hayat bilgisi ve sosyal bilgiler derslerinde çocuklar için felsefe: Bir eylem araştırması* (Thesis No. 627944) [Doctoral dissertation, Marmara University]. Council of Higher Education National Thesis Center.

Kılcan, B., & Akbaba, B. (2014). Sosyal bilgiler öğretim programında yer alan estetik değerine ilişkin öğrenci algılarının incelenmesi. *Journal of Theory and Practice in Education, 10*(4), 1047-1076.

Kostova, Z., & Radoynovska, B. (2010). Motivating students' learning using word association test and concept maps. *Bulgarian Journal of Science and Education Policy, 4*(1), 62-98.

Krippendorff, K. (2004). *Content analysis: An introduction to its methodology* (2nd ed.). Sage Publications.

Leckey, M. (2017). Guernica comes to school. Art, philosophy and life. In M. Gregory, J. Hayes, & K. Murris (Eds.), *The Routledge international handbook of philosophy for children* (pp. 137-144). Routledge.

Leddy, T. (2002). Aesthetics and children's picture-books. *The Journal of Aesthetic Education, 36*(4), 43-54. https://doi.org/10.2307/3301567

Lipman, M. (1978). *Suki*. Institute for the Advancement of Philosophy for Children.

Liptai, S. (2005). What is the meaning of this cup and that dead shark? Philosophical inquiry with objects and works of art and craft. *Childhood and Philosophy, 1*(2), 537-554. https://www.redalyc.org/pdf/5120/512051703011.pdf

Lone, J. M. (2012). *The philosophical child*. Rowman & Littlefield Publishers. https://doi.org/10.5040/9798881818777

Madeja, S. S., & Onuska, S. (1977). *Through the arts to the aesthetic: The CEMREL aesthetic education curriculum.* CEMREL, Inc.

Marquez, L. (2017). Philosophy in basic education: Towards the strengthening of the foundations of Philippine education. *Policy Futures in Education, 15*(1), 1-16. https://doi.org/10.1177/1478210317743650

McLeod, N., Stockell, K., Astbury, E., & McCann, J. (2020). Nurturing young children's self-identity and respect for difference through artwork as a provocation using philosophical enquiry: A pilot study. *Early Years, 42*(4-5), 480-495. https://doi.org/10.1080/09575146.2020.1817870

Mehdiyev, G., & Tozduman Yaralı, K. (2020). The effectiveness of training of philosophy with children on children's attitudes toward human values. *Journal of Bayburt Education Faculty, 15*(29), 251-279. https://doi.org/10.35675/befdergi.683349

Nutbrown, C. (2013). Conceptualising arts-based learning in the early years. *Research Papers in Education, 28*(2), 239-263. https://doi.org/10.1080/02671522.2011.580365

Nutbrown, C. (2020). *The arts and early childhood education.* The ORIM Blog. https://sheffield-real-project.sites.sheffield.ac.uk/orim-network/the-orim-blog/june-2020-the-arts-and-early-childhood-education

Okur, M. (2008). *Çocuklar için felsefe eğitimi programının altı yaş grubu çocuklarının, atılganlık, işbirliği ve kendini kontrol sosyal becerileri üzerindeki etkisi* (Thesis No. 226445) [Master's thesis, Marmara University]. Council of Higher Education National Thesis Center.

Oral, Ş. B. (2013). Can Deweyan pragmatist aesthetics provide a robust framework for the philosophy for children programme?. *Studies in Philosophy and Education,* (32), 361-377. https://doi.org/10.1007/s11217-012-9332-5

Özalp, H. K. (2020). Görsel sanatlar öğretmen adaylarının estetik alanına ilişkin bilişsel yapılarının kelime ilişkilendirme testi ile incelenmesi. *Erzincan Üniversitesi Eğitim Fakültesi Dergisi, 22*(2), 515-534. https://doi.org/10.17556/erziefd.730607

Pala, F. (2022). The effect of philosophy education for children (P4C) on students' conceptual achievement and critical thinking skills: A mixed method research. *Education Quarterly Reviews, 5*(3), 27-41. https://doi.org/10.31014/aior.1993.05.03.522

Parsons, M. J. (1994). Can children do aesthetics? A developmental account. *The Journal of Aesthetic Education, 28*(3), 33-45. https://doi.org/10.2307/3333399

Pekkarakaş, E. (2020). *Okul öncesi eğitim döneminde çocuklar için felsefe eğitiminin öğrencilerin yaratıcılık düzeylerine etkisi* (Thesis No. 648411) [Master's thesis, İzmir Demokrasi University]. Council of Higher Education National Thesis Center.

Read, H. (1931). *The meaning of art*. Penguin Books.

Rondhuis, N. (2005). *Philosophical talent: Empirical investigations into philosophical features of adolescents' discourse*. Veenman Publishing.

Santi, M. (2007). How students understand art: A change in children through philosophy. *Childhood and Philosophy, 3*(5), 19-33. https://www.epublicacoes.uerj.br/childhood/article/view/20507

Taş, I. (2017). *Çocuklar için felsefe eğitimi programı'nın 48-72 aylık çocukların zihin kuramı ve yaratıcılıklarına etkisi* (Thesis No. 460449) [Doctoral dissertation, Çukurova University]. Council of Higher Education National Thesis Center.

Türksoy, N. (2020). *Çocuklar için felsefe (P4C) eğitiminin ortaokul öğrencilerinin bilimsel sorgulamaya yönelik görüşlerine ve eleştirel düşünme becerilerinin gelişimine katkısı: Bir karma yöntem araştırması* (Thesis No. 640120) [Master's thesis, Alanya Alaaddin Keykubat University]. Council of Higher Education National Thesis Center.

Weber, R. P. (1990). *Basic content analysis.* Sage Publications. https://doi.org/10.4135/9781412983488

Wilson, R. A. (2010). Aesthetics and a sense of wonder. *Exchange: The Early Childhood Leaders' Magazine,* (193), 24-26. https://exchangepress.com/library/5019324.pdf

Education and Science

Original Article

Vol 51 (2026) No 226 379-403

A study on the readiness of schools for STEM education: are we really ready for it? *

Ümran Yiğit Dönmezoğulları [1], Gökhan Öztürk [2]

Abstract

To remain competitive in the rapidly evolving global economy, it is essential not only to utilize technology but also to produce it. Achieving this requires the development of skilled human capital across various disciplines. STEM education strengthens students' abilities to foster innovation by enhancing their critical and analytical thinking, and problem-solving skills, while also supporting affective development through increased empathy, ethical awareness, and social responsibility. This is a crucial step for economic growth, driven by the skilled workforce it helps create. However, successfully integrating STEM education into the current educational framework requires adequate preparation. The aim of this study was to assess the readiness of high schools in Ankara for STEM education. To achieve this, interview questions were developed based on the New York City Department of Education (NYCDOE) STEM framework, and interviews were conducted with teachers and administrators from four different types of schools. The schools' readiness was assessed through data analysis and further supported by qualitative themes identified from the interview responses.

Keywords

STEM
STEM education
STEM readiness
Teacher readiness
Student readiness

Article Info

Received: 08.31.2024
Accepted: 11.05.2025
Published Online: 04.30.2026

DOI: 10.15390/ES.2026.2513

Introduction

Over the last century, the world has witnessed remarkable advancements in science, economics, and technology. This has led to global competition among countries in areas such as the economy, technology, and defense. To succeed in these fields, nations require a scientifically informed population and a highly educated workforce in science, technology, engineering, and mathematics (STEM) (Chachashvili-Bolotin et al., 2016). Initially introduced and endorsed by the U.S. government in 2009, the STEM initiative aims to develop the human capital necessary for rapid economic growth and global competitiveness (Farina et al., 2016). For countries to remain competitive in the global market, they must adapt their education systems accordingly (Fensham, 2008). STEM education is key to this adaptation, as it enables students to apply their knowledge in practical, real-world contexts. When successfully integrated, STEM education can enhance students' creativity, problem-solving abilities, motivation, and

* This article is derived from Ümran Yiğit Dönmezoğulları's Master's thesis entitled "Exploring the STEM readiness level of high schools in Ankara", conducted under the supervision of Gökhan Öztürk.

[1] Middle East Technical University, Faculty of Education, Department of Mathematics and Science Education, Ankara, Türkiye, yigit.umran@metu.edu.tr

[2] Middle East Technical University, Faculty of Education, Department of Mathematics and Science Education, Ankara, Türkiye, gozturk@metu.edu.tr

overall academic performance. However, implementing STEM education in classrooms presents certain challenges that can affect a school's readiness to adopt this approach. A school is deemed ready for STEM education when its students, teachers, administrators, and physical infrastructure are adequately prepared to support STEM learning. Therefore, ensuring schools are prepared for STEM is a crucial step toward increasing the number of students who meet international STEM standards.

STEM Education

STEM, an acronym for Science, Technology, Engineering, and Mathematics, encompasses a multidisciplinary approach that goes beyond traditional lecture-based teaching strategies. Instead, STEM education emphasizes inquiry and project-based learning, which fosters critical thinking, creativity, innovation, and problem-solving skills among students (Breiner et al., 2012; Morrison, 2006). Beyond cognitive outcomes, STEM education also supports affective development, especially when implemented through real-world, socially relevant issues. In such contexts, students not only learn to solve problems but also develop empathy, ethical awareness, and a sense of social responsibility. For instance, Ozturk and Roehrig (2025) demonstrated that an integrated STEM unit centered on a socio-scientific issue significantly enhanced middle school students' abilities to take others' perspectives, recognize the societal role of science, and engage in nuanced emotional reasoning. Shim (2024) further conceptualized an "empathic STEM" framework, proposing that culturally responsive and human-centered STEM design tasks can foster empathy and ethical awareness among learners, especially in diverse classrooms. Similarly, Bush et al. (2024) argued that empathy-driven design thinking within STEM curricula not only promotes technical problem-solving but also nurtures students' sense of social justice and civic responsibility.

In order to drive advancements in technology and engineering, STEM education must strengthen students' understanding of scientific and mathematical concepts (Hernandez et al., 2014). Consequently, STEM education can be described as an interdisciplinary approach to learning that extends from early childhood through higher education, incorporating a wide range of educational activities (Gonzalez & Kuenzi, 2012). It is vital in today's educational context, as it offers numerous opportunities to connect theoretical knowledge with practical applications. It is a key element in helping nations maintain their competitiveness in the global economy. Focusing on technology and engineering from early childhood to future innovations, STEM education highlights the interconnectedness of science, mathematics, technology, and engineering, which cannot be treated as separate fields. Incorporating STEM education into the curriculum through methods such as inquiry-based learning, the use of digital tools, computer programming, and robotics helps students develop creativity, problem-solving abilities, critical thinking, motivation, and overall academic success.

Acquiring a comprehensive knowledge of STEM disciplines and the interdependence among them is a fundamental step toward attaining STEM literacy. STEM literacy means being able to understand and use ideas from science, technology, engineering, and math so that we can make sense of difficult problems and find better ways to solve them (Balka, 2011). This competence is essential for students to combine key knowledge, skills, and attitudes with the continuously changing technological context. Furthermore, it ensures that students are well-equipped upon graduation to face challenges in the technology-driven world (McDonald, 2016). Therefore, it is imperative for students to possess a deep understanding of the various disciplines in order to advance their careers in STEM.

The progress of modern society is heavily dependent on the integration of various scientific disciplines. Mathematics, in particular, plays a critical role in enabling scientific research, innovation, and technological advancement (Australian Academy of Science, 2006). While science provides students with an understanding of the natural world through the application of principles, concepts, and facts related to subjects such as physics, chemistry, and biology, it is not limited to these fields (National Research Council, 2012). With the help of technology and engineering, students can develop deeper connections between STEM concepts, thereby enhancing their creativity and higher-order cognitive skills. Additionally, engineering and technology practices provide a more comprehensive understanding of these concepts, ultimately enhancing students' motivation and success (Cunningham & Lachapelle, 2014). Moore et al. (2015) suggest that engaging in engineering practices enhances students' interest in STEM fields and careers by encouraging collaboration on problem-solving, fostering patience in managing processes, and utilizing 21st-century technologies to develop solutions.

STEM Education in Türkiye

Studies in Türkiye reveal that many young people are not particularly interested in science or creative thinking (Şahin et al., 2014). International assessments confirm this trend; Turkish students consistently score lower than their peers in other countries in science and math (Anıl et al., 2015; Oral & McGivney, 2013). The PISA National Report (Ministry of National Education [MoNE], 2016a) called for fundamental reforms in how STEM subjects (Science, Technology, Engineering, and Mathematics) are taught across all levels of education. To compete globally and foster innovation, Türkiye urgently needs a generation capable of creative thinking and problem-solving in STEM fields.

In response, both governmental and non-governmental institutions have taken steps to improve STEM education. The Ministry of National Education (MoNE, 2016b) emphasized that STEM is essential for students, teachers, and schools, proposing initiatives such as establishing STEM centers, conducting research, updating curricula, and training educators. Similarly, Turkish Industry and Business Association's (TUSIAD, 2017) report, "STEM Needs in Turkey by 2023," stressed the importance of swiftly forming a national STEM policy in collaboration with all stakeholders. Although the government had initial plans extending until 2019 (Altunel, 2018), recent years have witnessed continued efforts from private institutions, universities, and civil society. Importantly, STEM is not just about performing science experiments or learning to code; it is about equipping students with critical thinking and collaborative problem-solving skills (Bybee, 2013). Across Türkiye, a range of initiatives have emerged, including teacher training programs, STEM kits and public festivals that promote STEM awareness. The Ministry of National Education's (2016) report outlined concrete strategies that align with these activities (Arslan & Arastaman, 2021).

Studies evaluating these efforts suggest that they have begun to yield positive results. According to Çakır and Ozan (2018), STEM implementations in Türkiye have contributed meaningfully to student engagement and achievement in the targeted disciplines, supporting the broader goal of improving STEM education nationwide. In support of this trend, the PISA 2022 results revealed that Türkiye outperformed the average of all participating countries in Mathematics, Reading, and Science literacy (MoNE, 2024). The country has steadily improved in these domains for more than a decade, narrowing its performance gap with OECD countries. This positions Türkiye as one of the few nations to have shown consistent, long-term progress in the quality of education (MoNE, 2024).

To maintain and build on this progress, it is now essential to strengthen teachers' readiness, which plays a pivotal role in the sustained success of STEM education. Teachers need to develop a strong STEM-oriented mindset by moving away from teaching knowledge for its own sake and instead focusing on knowledge that can be applied in real-life context. This shift enables educators prepare students more effectively for real-world challenges by fostering practical application and innovation,

while also helping them develop critical thinking, creativity, and the essential skills needed for the future (Şanlı & Özerbaş, 2021).

Readiness

Readiness is a key factor in successfully completing tasks, as it requires the necessary conditions to achieve a specific goal. Merriam-Webster's Collegiate Dictionary (Merriam-Webster, 1999) defines readiness as "a state of preparation," though this definition is broad and context dependent. Hersey and Blanchard (1988) highlighted the importance of performing tasks willingly when discussing readiness. In the context of STEM education, it is essential to consider the preparedness of both teachers and students to ensure effective implementation. Lynch and Smith (2016) state that when members of a school are prepared to manage change, it signifies their readiness for improvement, underscoring the significance of teacher and student preparedness in achieving success in the evolving context of Science, Technology, Engineering, and Mathematics (STEM) education. In this sense, readiness includes more than just having the right tools or basic knowledge; it also requires motivation, confidence, and openness to new ways of teaching and learning. Hall and Hord (2011) further argue that lasting changes in education depend on support from school leaders, collaboration among teachers, and access to professional development. They also note that change does not occur instantly but is a gradual process through which both individuals and institutions gain understanding, develop skills, and become confident in applying new methods.

Teacher Readiness

> "I have come to a frightening conclusion. I am the decisive element in the classroom. It is my personal approach that creates the climate." (Ginott, 1976).

The quote by Ginott (1976) serves as a powerful reminder that teachers play a decisive role in setting the tone of the classroom. Their personal approach and daily mood are key factors that shape the classroom climate and, ultimately, the overall learning experience of their students. The relationship between teachers and students is complex and important, as both are integral to the learning community. Creating a positive and effective learning environment is crucial, and it all begins with the teacher's readiness. In particular, teachers who are open-minded are more likely to create environments that fosters creative thinking, encouraging students to become risk-takers and open-minded individuals (Crane, 2017). Motivation is essential to the teaching process, as it empowers teachers to share knowledge and work interdisciplinarily. Motivated teachers are energized, eager to learn and teach, and utilize their full potential to go the extra mile (Sinclair, 2008).

Teacher readiness becomes even more critical due to the interdisciplinary nature of the content and the need for inquiry-based and student-centered pedagogies. Research shows that even content-proficient teachers may feel unprepared to implement STEM effectively without targeted support (Margot & Kettler, 2019). Moreover, many educators express uncertainty when it comes to integrating engineering principles, using technology tools, or designing cross-curricular projects. Therefore, readiness also includes pedagogical adaptability, technological fluency, and a growth mindset for learning and teaching in rapidly evolving fields. Similarly, Darling-Hammond (2020), emphasizes that teachers feel more prepared when they receive strong support and opportunities for continuous learning, including professional development, peer collaboration, and leadership that values experimentation and risk-taking. As Hall and Hord (2011) argue, readiness is not an instant condition but a gradual process through which teachers build readiness over time by engaging with new methods, reflecting on their practice, and receiving meaningful feedback.

In conclusion, a teacher's readiness forms the foundation for creating a positive and effective learning environment. By maintaining an uncluttered mind, being open-minded, motivated, willing, and capable, teachers can inspire and empower their students to achieve their full potential. In STEM education, teacher readiness acts as a bridge between curriculum goals and actual student engagement. A well-prepared teacher doesn't merely deliver content; they model curiosity, resilience, and the lifelong learning mindset that STEM demands.

Student Readiness

The readiness of both students and teachers is crucial in the learning process, as they mutually influence one another. According to Schunk (2012), readiness refers to what students are able to learn or accomplish at various developmental stages. Factors such as prior knowledge, interests, attitudes, and abilities shape a students' levels of readiness (Çetin, 2016). Learning is described by Arozaq et al. (2017) as a process of behavioral change, and the duration of this process is influenced by the students' readiness (Harman & Çelikler, 2012). When students exhibit a high level of readiness, they tend to perform successfully in activities and experience greater overall satisfaction. Thorndike's Law (Harman & Çelikler, 2012) outlines the following principles:

- When individuals are ready to engage in an activity, they derive satisfaction from participation.
- If individuals are prepared to perform an activity but are not permitted to do so, it may result in frustration.
- Forcing individuals who are unprepared to participate in an activity can lead to anger.

Assessing readiness levels for STEM education helps teachers align their preparation with the needs of their students. Student readiness extends beyond academic ability; it encompassed problem-solving disposition, collaborative skills, and comfort with ambiguity, all of which are essential in inquiry-based learning environments. According to Bybee (2010), effective STEM instruction depends on students' capacity to engage in authentic problem-solving tasks that require both prior knowledge and a willingness to experiment. Moreover, research shows that students' interest in STEM careers is closely linked to their self-efficacy and early exposure to integrated STEM experiences (Radloff & Guzey, 2016). However, without foundational confidence in disciplines such as mathematics or science, they may hesitate to participate actively in cross-disciplinary activities. As Le et al. (2023) point out, even when students are curious about STEM, gaps in fundamental understanding, particularly in mathematics, can limit their engagement. Therefore, teachers must assess not only content knowledge but also emotional and motivational readiness indicators such as persistence, curiosity, and risk tolerance.

Finally, readiness should be understood as a dynami construct rather than fixed state, as it is influenced by classroom climate, peer interaction, and instructional design. A supportive and responsive environment can enhance even initially unprepared students' readiness over time. Teachers who differentiate instruction based on readiness are better positioned to scaffold learning and foster success among diverse learners in STEM classrooms.

NYC STEM Framework

The NYC STEM Education Framework serves as an effective tool that offers a structured approach for schools seeking to enhance their implementation of STEM initiatives. The framework aims to transform teaching and learning by shifting from multidisciplinary and interdisciplinary approach to a transdisciplinary one, thereby creating a more enriching experience for both students and educators. STEM-educated students develop crucial skills such as critical and interdisciplinary thinking, inquiry-based and real-life learning, and participation in transdisciplinary learning processes (Altunel, 2018).

The study identified several key criteria for assessing schools' readiness levels, including socio-economic disparities, exam-oriented education systems, and teacher training. The framework employed in the research served as an effective analytical tool, providing valuable insights to the study. However, it is important to note that this framework is not designed for direct evaluation of schools; rather, it aims to identify their preparation needs. In this context, readiness refers to a school's capacity to effectively implement STEM education. The framework encompasses various subcategories, ranging from establishing a STEM vision within schools to supporting career planning in related fields. It comprises four main domains and twelve indicators (Table 1), which define the readiness levels categorized as "Early," "Emerging," "Integrated," and "Fully Integrated." A detailed explanation of all domains and indicators is presented below.

Domain 1: School Vision and Structures for Success

The STEM mission and vision of the school must be clearly articulated and effectively communicated to all stakeholders, while the school's culture, program evaluation processes, and budgetary resources should be aligned to support these goals. The STEM Mission and Vision indicator helps assess how well the school's staff integrates and embraces the STEM mission and vision, with evidence of commitment reflected in educational plans that incorporating STEM, meetings focused on integration, and professional development opportunities related to STEM. Complementing this, the STEM-centric Culture indicator evaluates the extent to which the school's environment supports STEM learning. A positive culture that encourages innovation, risk-taking, and interdisciplinary collaboration among all members is vital for achieving full integration. Moreover, the acceptance of student-centered inquiry, engineering practices, digital literacy, and project-based learning by all stakeholders, together with access to adequate STEM classes and laboratories, represents additional critical elements for success.

The evaluation of the degree to which STEM education has been implemented is captured within the STEM Program Evaluation indicator. Achieving full integration critically depends on the availability of STEM-rich resources and a leadership team that involves families, school administrators, and other key stakeholders. Another important indicator, Budget and Resource Management, addresses the adequacy of the school's STEM budget and the allocation of funds for related initiatives. Maintaining accurate budget records and having a dedicated team to oversee these finances are essential. However, limited access to technology and other essential resources for STEM education may hinder effective integration.

Domain 2: Curriculum, Instruction, and Assessment

This domain emphasizes a transdisciplinary instructional approach that fosters student-centered inquiry, project-based learning, and engineering practices within a STEM-focused curriculum, assessing instructional quality, authentic assessments practices, and staff capacity as key indicators of success. The Academic Rigor and Instructional Quality indicator underscores the necessity of a clearly defined STEM program that promotes critical thinking, real-world problem-solving, and active learning through the application of 21st-century skills. Integrating STEM applications into curricula, lessons, and units, along with adapting teaching practices to align with the school's core mission and vision across all STEM courses, is essential for enhancing integration. The STEM-centric Curriculum indicator further emphasizes the importance of a robust STEM education program that cultivates innovative and critical thinking, engineering design, and scientific literacy. Providing STEM opportunities both inside and outside the classroom enables students to apply STEM concepts in real-world contexts, while dedicating sufficient time to support STEM education further enhances integration.

To evaluate success, the Authentic Assessments indicator emphasizes the effective implementation of authentic assessment tools, such as projects, portfolios, and presentations. The regular use of learning cycles, together with formative and summative assessments, provides students with opportunities for self-assessment and targeted feedback, thereby enhancing integration. The Staff Capacity indicator focuses on improving professional learning in STEM content areas and strengthening teachers' pedagogical expertise. Additionally, the development of an effective collaborative plan among

STEM stakeholders, along with the dissemination of valuable information about STEM opportunities and programs, represents a crucial element of this domain.

Domain 3: Strategic Partnerships

This domain focuses on STEM partnerships that engage families, universities, community organizations, and businesses. The STEM Partnerships indicator assesses the effectiveness of intentional collaborations in STEM education that involve universities, families, local organizations, and businesses. Successful integration of STEM education depends on collaboration between STEM partners and educators, effective communication with families, and the systematic organization of STEM-related events. Additionally, collaboration with other schools that share similar interests in STEM can further enhance the overall level of integration.

Domain 4: College and Career Readiness

This domain primarily aims to guide students toward a STEM-oriented career trajectory, beginning in elementary school and continuing through high school. The STEM Pathway Preparation for Elementary School indicator underscores the importance of utilizing resources such as College Talk to foster early awareness of STEM-related careers. To ensure full integration, it is vital to demonstrate evidence of opportunities both inside and outside the school, including guidance on STEM careers for parents, students, and teachers.

The Access to STEM College and Career Opportunities for Middle and High School Students indicator differs from the previous one, focusing specifically on STEM-related college and career opportunities for middle and high school students. Providing guidance to students and families on essential factors such as self-discipline, homework, and GPA is crucial for enhancing career readiness and increasing integration levels. Similarly, the Planning Student Outreach and Support for Pre-K-12 STEM Initiatives indicator highlights the intentional support schools should offer to students and families regarding future STEM careers and related pathways, underscoring the importance of proactive engagement across all grade levels.

Table 1. The domains and indicators of STEM Framework

DOMAINS		**INDICATORS**
Domain 1	School Vision and Structures for Success	*STEM Mission and Vision *STEM-centric Culture *STEM Program Evaluation *Budget/Management of Resources
Domain 2	Curriculum, Instruction and Assessment	*Academic Rigor and Instructional Quality *STEM-centric Curriculum *Authentic Assessments * Staff Capacity
Domain 3	Strategic Partnerships	* STEM Partnership
Domain 4	College and Career Readiness	* Access to STEM College and Career Opportunities for Middle and High School Students *Planning Student Outreach and Support for Pre-K-12 STEM Initiatives

Method

Research Design

The instrumental case study design was deemed appropriate for this qualitative research. This method allows researchers to examine a specific case, selected for its relevance to the research question, to gain valuable insights or to serve as a tool for testing or illustrating theoretical concepts. Thus, an instrumental case study focuses on a particular problem rather than on the case itself, using the situation to gain a deeper understanding of the issue (Stake, 1995).

In this study, the primary focus is the readiness of high schools, and the inclusion of different school types constitutes multiple cases. Therefore, a multiple-case study approach was adopted. According to Baxter and Jack (2008), this approach is effective for exploring similarities and differences among the cases. Moreover, findings from multiple-case studies tend to be robust and credible (Baxter & Jack, 2008). Engaging in multiple-case research allows for a more in-depth exploration of research problems and contributes to theoretical development (Eisenhardt & Graebner, 2007).

Sample

The high schools in Ankara serve as the cases for this study, focusing on their respective levels of STEM readiness. Four distinct schools were selected for this purpose: one science high school, one private high school, one vocational high school, and one regular public high school. The selection process was guided by a combination of accessibility considerations and purposeful sampling criteria. Initially, the researcher contacted several schools in Ankara with which academic or professional relationships already existed, facilitating access and communication. Although a few schools declined to participate due to administrative workload or scheduling conflicts, the final sample comprised those that responded positively and demonstrated a willingness to participate in the study. The inclusion of four distinct school types was intentional and aimed to capture diverse STEM readiness contexts based on institutional resources, curricular orientation, and student demographics. Practical considerations such as ease of access, variation across school types, and availability of key staff members (e.g., science teachers and deputy principals) also played a role in the final school selection.

Furthermore, the study adopted a multiple-case design in which one school from each type (science, private, vocational, and regular public high schools) was selected. This approach ensured diversity across cases and supported analytic generalization, while maintaining the practicality of the study's scope for detailed analysis. As Baxter and Jack (2008) explain, a multiple-case study allows the researcher to conduct both within-case and cross-case analyses, thereby identifying similarities and differences among the cases. For this reason, including one school from each type was deemed sufficient to ensure variation and enable meaningful comparisons, while keeping the study manageable for in-depth analysis.

The private high school distinguished itself by offering an international curriculum alongside the national one. This school had a diverse student body, including international students and staff, and was equipped with well-maintained laboratories. Thematic classes, such as geography, were commonly offered. While the school had three fully equipped science laboratories, it lacked a computer laboratory.

The science high school, selected based on its performance in the Ministry of Education's high school entrance examination, is recognized as one of the top high schools in both the city and the country. Although structurally similar to the private high school, its laboratories and equipment were comparatively older. This school had two physics laboratories, two chemistry laboratories, and two biology laboratories, all providing sufficient infrastructure for students to carry out projects and experiments. The facility included 16 classrooms, each equipped with an interactive whiteboard and internet access. Notably, this school actively used 3D printers, a feature not found in the other selected schools. Additionally, there were six workshop classrooms available for instructional and extracurricular activities.

Vocational high schools aim to contribute to the national workforce by educating students in technical and applied fields. The vocational high school included extensive facilities and laboratories. There were twelve laboratory and workshop classrooms in total. The laboratory resources were actively utilized. Students who did not pass the national entrance examinations and opted for alternative school types attended regular public high schools. The curriculum at such institution follows the national standard. In the regular public high school, classrooms were equipped with smartboards; however, they were overcrowded, and no laboratories were in use. It was completely exam-oriented school, with a primary focus on university entrance preparation

Given that participants were easily accessible, geographically close, and willing to participate, the sampling method employed can be described as convenience sampling, which is non-random in nature as noted in Dörnyei's (2007) work on research methods. Therefore, purposive sampling was also employed in this study. This approach emphasizes the concept of saturation (Miles & Huberman, 1994), which refers to continuing data collection until no new information emerges. Accordingly, no predetermined sample size was set at the outset of the study. Ultimately, at least one deputy principal or school principal, along with two teachers from each school, participated. In total, nine science or mathematics teachers and four deputy principals or school heads participated in the study. The selection of participants followed a homogeneous sampling strategy, and all interviews were conducted voluntarily.

Data Collection and Analysis

The STEM Framework Criteria developed by the New York City Department of Education (NYCDOE) were used in this study to guide construct interview questions. The framework comprises four domains, each containing indicators and corresponding criteria. The indicators within these domains guided the selection of interview questions, and schools' readiness levels were categorized as Early, Emerging, Integrated, or Fully Integrated. To collect meaningful and relevant data, semi-structured interviews were conducted on a voluntary basis. There was no fixed time limit for these interviews, which lasted approximately twenty minutes on average. The interview questions were carefully designed in alignment with the framework, to ensure that the data gathered were purposeful and aligned with the study's objectives.

To ensure the validity and reliability of the interview questions, expert feedback was obtained during the development process. Specifically, two STEM education scholars and one qualitative research specialist were consulted. These experts reviewed the questions for clarity, relevance, and alignment with the NYC STEM Framework, and their feedback was incorporated into the final version. Additionally, the questions were piloted at the researcher's school to assess their clarity, comprehensibility, and flow. Based on pilot results and expert feedback, minor revisions were made to refine the wording and sequencing. The final version of the instrument was approved by the study supervisor, contributing to its content and face validity.

During the analysis phase of the study, all interview audio recordings were transcribed verbatim to ensure accuracy and depth of the data. This included detailed accounts from two science and mathematics teachers and one administrator from each participating school. The analytic process was further supported by memos that the researchers documented before, during, and after the interviews. These memos facilitated the integration of codes and aided in the formulation of interpretive judgements, which is one of the key strengths of qualitative research (Denzin & Lincoln, 2011). Following transcription, the researchers conducted a pre-coding stage by reviewing the transcripts and highlighting relevant segments. The interview questions were structured around the NYC STEM framework, and the responses from both experts and practitioners were analyzed using the Constant Comparative Method (CCM). This method allows for continual comparison between collected data and emerging categories, making it particularly suitable for the objectives of this study (Creswell & Poth, 2016).

The analysis began with open coding, which involved examining key words, phrases, and statements from the interview data. Following Kolb's (2012) recommendation, the researchers compared the data by questioning what was understood and what remained unclear. In the second phase of the analysis, the data was reorganized by establishing connections between categories, ultimately refining the codes to align with the main themes. As a result of this analysis, several themes emerged from practitioners' perspectives regarding the schools' levels of STEM readiness.

In a later stage, primary headings were established under each domain, resulting in a total of twenty-five headings. Each heading contained four distinct indicators, which were scored on a scale from 1 to 4. For Domain 1, three main headings were identified, with four indicators associated with each. Each school was categorized under an indicator based on the outcomes of the interview questions. A summative approach was employed in this section to identify commonly used words and assess their frequency (Hsieh & Shannon, 2005). Although determining the frequency of recurring words may suggest a quantitative analysis (Kondracki et al., 2002), the primary focus remained on content interpretation, thereby characterizing it as a qualitative study (Hsieh & Shannon, 2005).

The four-point scoring system (1 = Early, 2 = Emerging, 3 = Integrated, 4 = Fully Integrated) was directly derived from the NYC STEM Framework, which categorizes school readiness into four levels across each subdomain. Each subdomain in the framework includes observable indicators aligned with one of these readiness levels. During data analysis, participant responses were examined and matched to these indicators based on content analysis and frequency of indicator-related expressions. When responses clearly reflected the characteristics of a particular readiness level, for instance, if a school lacked a STEM mission or practices, it was coded as "1 = Early"; if STEM activities were occasional and fragmented, it was coded as "2 = Emerging"; if STEM was regularly implemented in a structured way, it was coded as "3 = Integrated"; and if it was deeply embedded across the curriculum and school culture, it was coded as "4 = Fully Integrated." This process was consistently applied across all domains and the four selected schools. At the conclusion of the scoring phase, each school's total readiness score, ranging from a minimum of 25 to a maximum of 100, was used to classify its STEM readiness level. Based on this structure, schools scoring between 25 and 43 were classified as Early, those scoring between 44 and 62 as Emerging, scores between 63 and 81 were categorized as Integrated, and scores from 82 to 100 were labeled as Fully Integrated. These intervals were designed to reflect an even distribution across the range of possible scores and to align conceptually with the four-tier framework established by the NYC STEM criteria.

To ensure the reliability of the analysis, the coding process was conducted collaboratively by the primary researcher and two field experts. Each coder independently reviewed and scored the transcribed data. The results were then compared across coders to evaluate inter-coder agreement. Any discrepancies in scoring were discussed and resolved through consensus. This triangulated coding approach enhanced the trustworthiness and consistency of the final scoring across all four domains, thus strengthening the study's methodological rigor.

Results

Domain Scores

The domain scores are presented in the figures below, with Domain 1 focusing on the STEM vision of the schools and the structures necessary for success. This domain comprises four subdomains that collectively assess how well each school articulates its STEM vision and the supporting frameworks in place to achieve it. There were three subcategories under Domain 1.1, which dealt with the goal and vision of STEM. Participants received scores between 4 and 12 from this domain because there was a minimum score of 1 and a maximum score of 4 for each category. Domain 1.2 was about STEM-centric culture and included 4 categories and the scores of participants ranged from 4 to 16. STEM program evaluation was covered in Domain 1.3, and STEM budget was covered in Domain 1.4. The greatest score for these domains is 8, and the lowest value is 4. The highest possible score for participants in this area has been set at 36 out of 100 (Figure 1). Domain 2 included four subdomains and was primarily

concerned with STEM curriculum, instruction, and assessment. Following analyses using the same methodology as previously, the maximum score for this domain was 44 out of 100 (Figure 2). The same process was used for Domains 3 and 4, with a maximum score of 8 and 12 respectively out of 100 (Figure 3).

Results of Domain 1 (School Vision and Structures for Success)

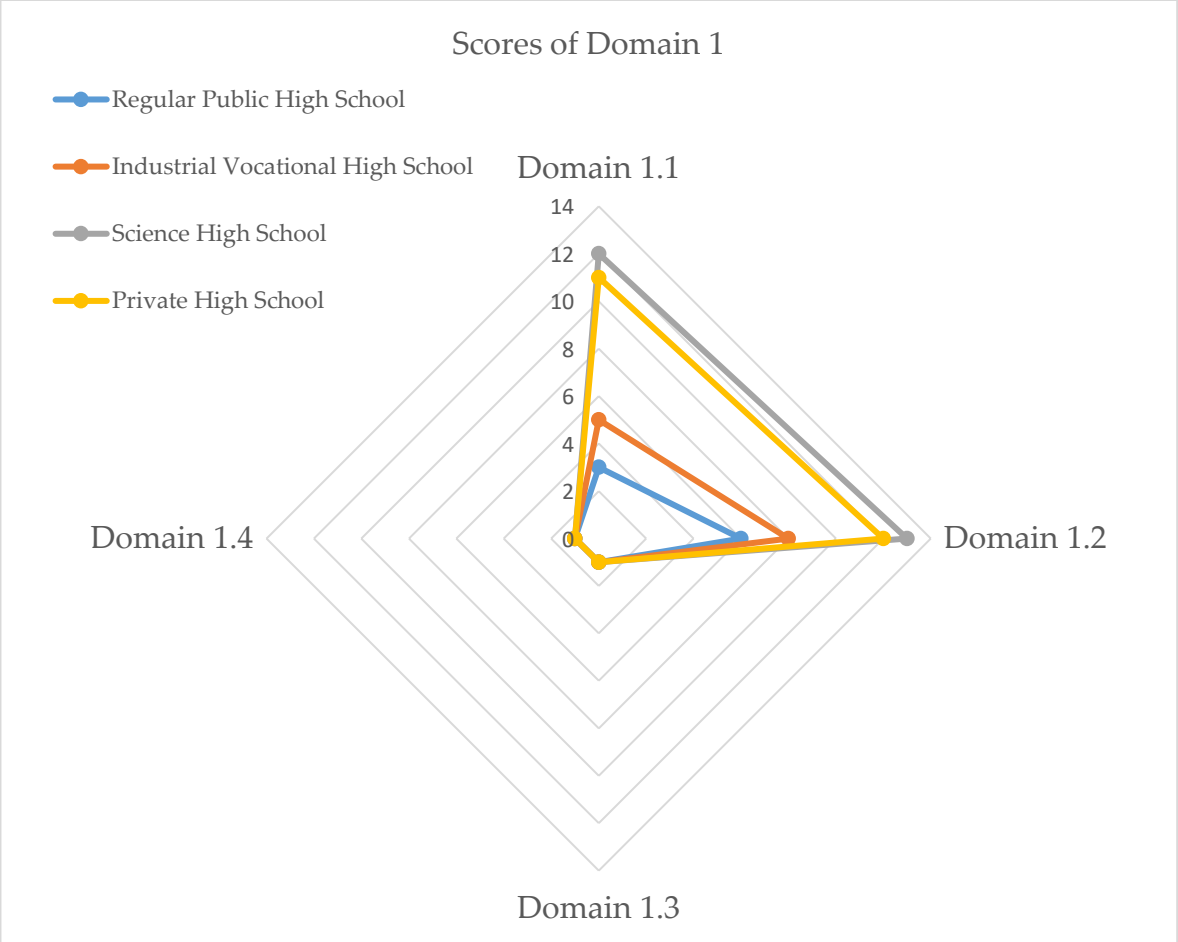


Figure 1. Scores of Domain 1

Domain 1.1 (STEM Mission and Vision): The data collected from the participants indicated that none of the individuals at the regular public high school were familiar with STEM or expressed any interest in it. In the vocational high school, only a few participants had heard the term but were uncertain about its meaning. Among those who were aware, only a small number expressed interest in STEM education. Some participants believed that STEM referred to an organization established by various institutions, while others correctly identified it as an acronym for science, technology, engineering, and mathematics, or associated it with these disciplines.

Conversely, all participants from both the science high school and the private high school were familiar with STEM education and expressed interest in it. Participants from the private high school described STEM as an "interdisciplinary approach." Consequently, while the regular public high school and vocational high school had not yet established a STEM mission and vision, the science high school had already defined its mission and vision, and the private high school was on the verge of fully implementing them.

The results indicate that participants generally experience confusion regarding the definitions of STEM and STEM education. Teachers at the private high school possess a comprehensive understanding of STEM due to their involvement in the International Baccalaureate (IB) Programme, which emphasizes critical thinking, self-reflection, and interdisciplinary connections. Similarly, participants from the science high school demonstrate a strong understanding of STEM, likely owing to their collaboration with research universities. In light of these findings, it is essential for schools to establish a clear and coherent STEM mission and vision.

In contrast, participants from the other schools demonstrated little interest in STEM, primarily due to the focus on university entrance examinations and concern sregarding students' academic performance. The perceived low academic level was identified as a barrier to providing high-quality science, technology, engineering, and mathematics (STEM) education. However, this should not serve as an excuse, as STEM activities can be adapted to suit all educational levels and implemented at any stage of schooling (U.S. Department of Education, 2015). Therefore, although STEM is not currently a priority for these schools, its future integration remains uncertain.

Domain 1.2 (STEM-centric Culture): When examining the STEM-centered Culture domain, it becomes evident that the regular public high school and the vocational high school offer only a limited number of STEM-related courses and interdisciplinary teaching opportunities. Although these schools have laboratories, they lack dedicated classes and facilities specifically designed for STEM activities. Participants from the regular public high school noted that overcrowded classrooms hinder the effective implementation of STEM education.

In contrast, the private high school and the science high school provide a more favorable environment for STEM integration. These institutions offer a greater number of STEM-oriented programs, dedicated STEM classes, and laboratories that facilitate interdisciplinary courses. Furthermore, both schools employ a larger number of staff members who help foster a supportive atmosphere for STEM education, although this remains insufficient to fully realize the potential of STEM initiatives. Overall, the findings indicate a clear disparity in STEM readiness and culture among these schools, with the science and private high schools leading in terms of available resources and institutional support.

Although the schools may not currently have a STEM-centered culture, some are willing to take the first steps. While certain clubs at the science high school and the private high school focus on STEM, this is not a school-wide initiative. Given that the IB Programme includes STEM projects, attending a private school appears to be advantageous. One positive takeaway from the findings is that participants from all schools now understand that incorporating technology into the curriculum does not equate to STEM. At this point, cultivating a STEM-centric culture requires increased awareness. Teachers who use technology in their lessons and conduct science experiments in their courses believe they are practicing STEM, but in reality, they are not. Another crucial element in developing a STEM-centric culture is establishing the right conditions and gaining the support of administrators and teachers, although some schools may find this challenging. For example, teachers in regular public high schools are not motivated to participate in STEM activities. This may be due to the exam-oriented Turkish education system or to individual factors, but this is a topic for another study. It appears that schools still need time and motivation to cultivate a genuine STEM culture.

Domain 1.3-1.4 (STEM Program Evaluation-Budget/Management of Resources): Based on the data obtained from the STEM Program Evaluation and STEM Budget domains, the evaluation process of STEM education in the schools was found to lack any evidence. Furthermore, it was reported that no budget had been allocated for STEM education. As a result, all schools scored 1 out of 4 in these domains. The participants' responses indicate that STEM education is not currently being implemented in schools; hence, the concept of a STEM budget is virtually non-existent. These schools are struggling to meet even their most urgent financial needs, let alone allocate funds for STEM education. Consequently, the STEM evaluation process is also not being considered. Schools must first cover their

daily expenses, such as project costs and laboratory supplies, before they can think about investing in STEM education. Almost all of the schools that were contacted reported lacking an adequate budget. In times of necessity, they manage to secure funds through donations and parental contributions.

Results of Domain 2 (Curriculum, Instruction, and Assessment)

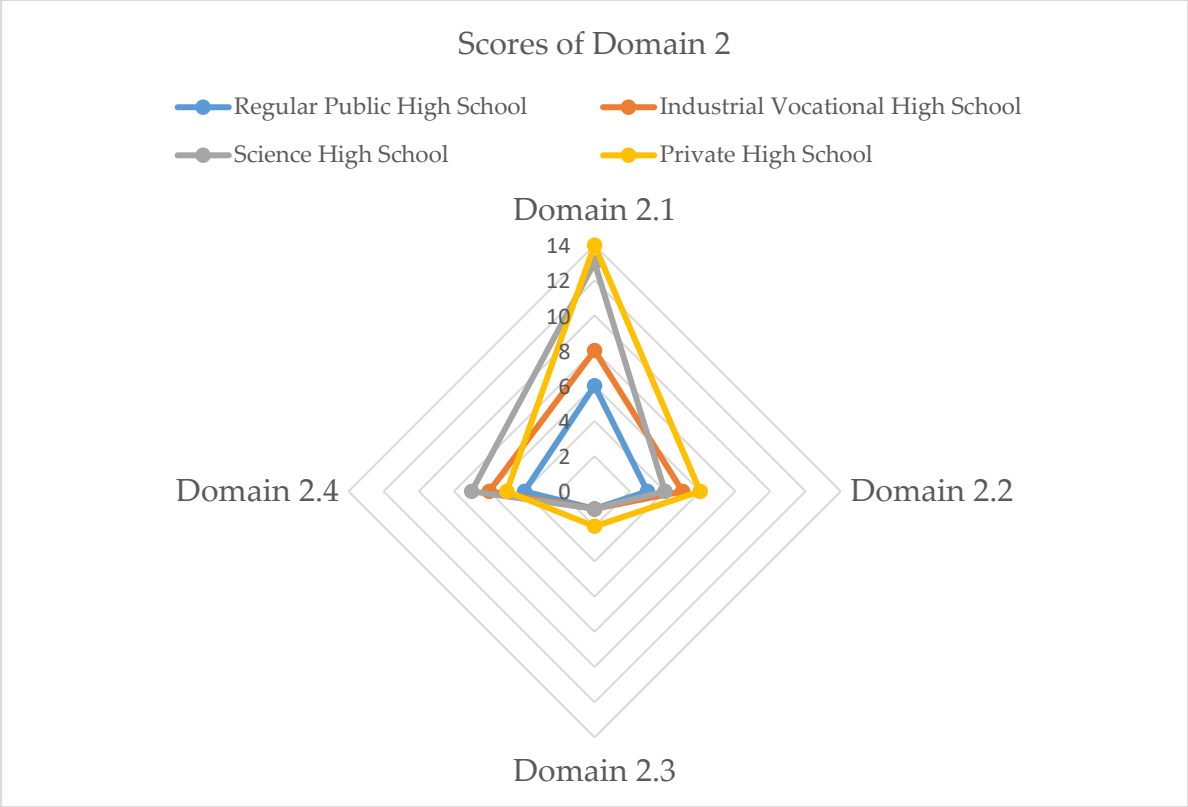


Figure 2. Scores of Domain 2

Domain 2.1 (Academic Rigor and Instructional Quality): Drawing on the insights obtained from the Academic Rigor and Instructional Quality domains, this section identifies four subdomains. The science high school and the private high school frequently incorporate student-centered courses and inquiry-based methods into their teaching, whereas such practices are rare in other institutions. Moreover, the regular public high school lacks STEM applications, and the teaching strategies employed there are not suitable for STEM instruction. Currently, none of the schools fully implement STEM methodologies, and neither of them has a STEM curriculum that is completely aligned with the instructional strategies in use.

Inquiry-based learning, the integration of STEM into classroom instruction, and the use of STEM applications often rely heavily on the individual initiative of educators within the school system. This dependency is particularly evident in private high schools, where teachers tend to be more cautious. The reason lies in the structure of the IB Programme, which formally requires inquiry-based methods and student-centered pedagogies, thereby placing additional demands on teaching practices. Teachers at the science high school expressed a willingness to integrate STEM into their lessons; however, the university entrance examination poses a significant barrier, as 10th-grade students often prioritize exam preparation over exploring additional subjects. Conversely, teachers at the regular public high school showed little interest in implementing STEM education. Overall, the teaching methods employed are not conducive to STEM learning, primarily because Türkiye's national curriculum does not explicitly promote activities directly related to STEM. Additionally, teachers voiced concerns about the low academic levels of students and the limited time available for instruction.

The interviews revealed that some teachers were unfamiliar with the inquiry method, highlighting a more urgent need for teacher education than for STEM education in the country. As Vermunt (2014) points out, the processes of student learning and teacher learning are interconnected; the quality of teacher learning directly influences the quality of student learning. Therefore, it is essential to recognize that teacher learning, enabling educators to acquire new skills, is an ongoing process (Solheim et al., 2018).

Domain 2.2 (STEM-centric Curriculum): This domain examines the extent of STEM integration within courses, the existence of a STEM content curriculum, and the opportunities available for applying STEM concepts. While the regular public high school never integrated STEM subjects, the other schools rarely did. Furthermore, no school had a dedicated STEM curriculum. STEM was only partially incorporated into the private high school curriculum. Although the participating schools generally did not provide STEM applications for their students, they noted that the vocational high school and the private high school occasionally did. However, participants from the vocational high school reported that their students regularly went on field trips to manufacturing facilities and factories utilizing new technologies. Although the participants thought these trips allowed students to observe real-world applications of STEM fields, the curriculum did not include specific objectives linking the field trips to STEM education.

Domain 2.3 (Authentic Assessment): Based on the information gathered from the participants, authentic assessment was not used in any institution other than the private high school. As a result, the private high school received a score of 2 out of 4, whereas the other schools scored 1 out of 4. Authentic assessment enables students to apply their foundational knowledge and skills to real-world contexts, distinguishing it from traditional assessment methods (Mueller, 2008). Unfortunately, it must be acknowledged that the schools continue to employ outdated evaluation techniques such as multiple-choice exams, fill-in-the-blank items, true-false questions, and matching exercises. Once again, the IB Programme is the only one with policies that support the use of authentic assessment, albeit ineffectively.

Domain 2.4 (Staff capacity): An analysis of the staff capacity in this domain revealed that not all schools had teachers with the necessary experience or pedagogical understanding of STEM subjects. Furthermore, while some schools engaged in collaboration occasionally, neither the private high school nor the regular public high school collaborated on STEM-related issues. Additionally, the science high school provided teachers with the most frequent unplanned STEM support, whereas the regular public high school offered none at all.

Based on participant data, it was reported that teachers do not feel they possess sufficient academic knowledge of STEM education. Teachers at vocational high schools are open to teaching STEM subjects. Additionally, they stated that while administrators can provide some support in this area, it is insufficient on its own, and schools' physical environments must be adapted to accommodate STEM. However, there are no formal expectations or requirements for educators or administrators to develop their STEM competencies, and there is no organized dissemination of STEM knowledge within classrooms. Participants rely on their own motivation and individual efforts to sustain this process.

Results of Domain 3 (Strategic Partnerships)

Based on the information gathered from the participants, the schools lacked strategic STEM partnerships with universities or companies, with the exception of the vocational high school and the science high school. Additionally, there was minimal collaboration with families on STEM initiatives and limited partnerships with other schools. Participants from the regular public high school indicated that parents' main concern is their children's performance on exams rather than their long-term futures, resulting in no educational push for STEM at the school due to a lack of family demand. Nonetheless, partnerships in education with families, universities, and colleges are essential for the advancement of STEM education.

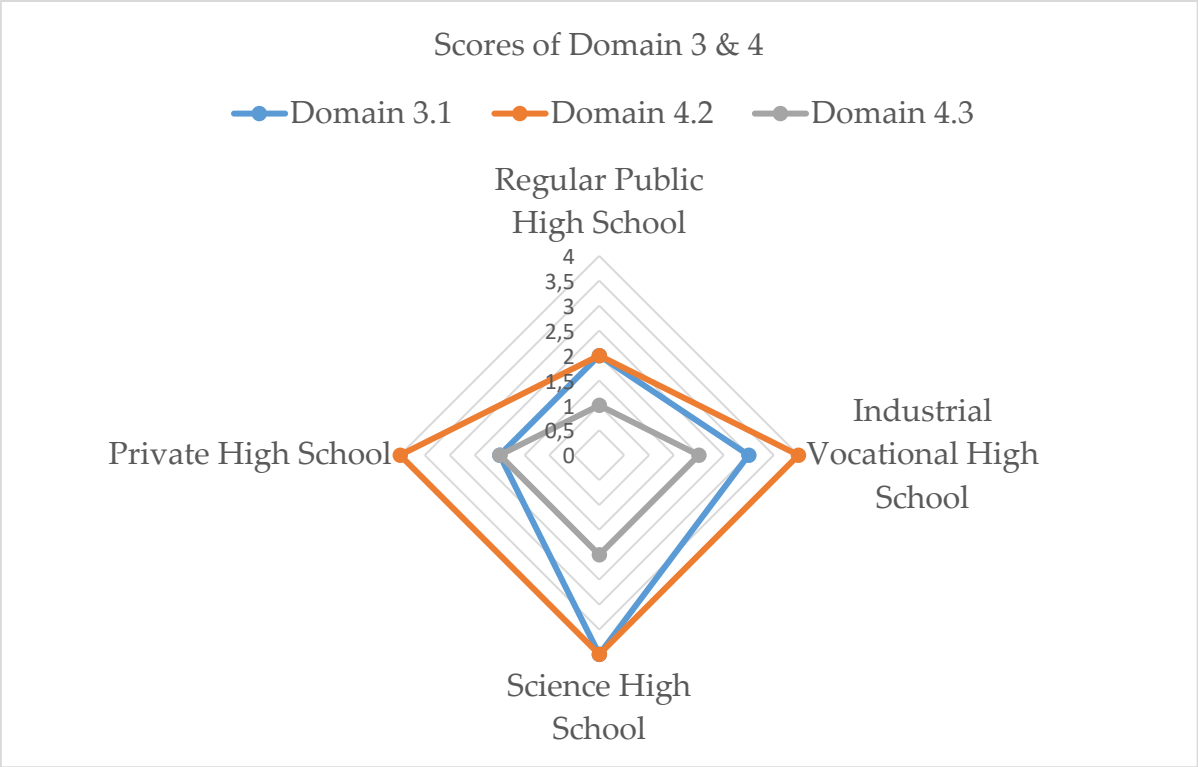


Figure 3. Scores of Domain 3&4

Results of Domain 4 (College and Career Readiness)

The findings of the Planning Student Outreach and Access to STEM College and the Career Opportunities domains indicate that there were no resources available at the regular public high school and that only three schools offered opportunities for students to learn about STEM careers. In addition to the lack of opportunities, very few students were aware of STEM-related careers, and not a single student at the regular public high school knew about them. Lastly, although such initiatives were infrequent in other schools, students at the regular public high school were unable to receive support for future planning in STEM-related areas. Since the primary goal of the nation's education system is to prepare students for university entrance examinations, families and students tend to focus solely on academic achievement. Therefore, it is not surprising that career planning for STEM is either absent or severely limited within schools.

In the concluding section of the analysis, key statements were extracted from participants' responses. Subsequently, meaning units were developed from these significant statements, and these units were organized into broader themes. The themes identified in this study include the need for greater knowledge and awareness of STEM, the necessity of a STEM-integrated curriculum, the impact of university entrance exam anxiety as a barrier to STEM education, and the need for adequate budgets and resources to support STEM initiatives.

Theme 1: The need for knowledge and awareness of STEM

Based on the interview findings, it is evident that a significant number of participants have limited knowledge of STEM, with some possessing incomplete or inaccurate information. As a result, overall STEM awareness among participants remains considerably low. While some recognize STEM as a term, others perceive it as a process, a resource, or a novel concept. Therefore, it is crucial to establish a clear and consistent definition of STEM to ensure that educational institutions are adequately prepared to deliver STEM-oriented instruction.

Here are some quotes that reflect different perspectives on the definition of STEM:

"I heard of STEM, but I haven't heard of STEM education." (T)

"STEM is something that many organizations do." (T)

"STEM is a learning method in the fields of science, technology, engineering and mathematics." (A)

"STEM means being interested in mathematics, physics, chemistry, biology and technology." (T)

"STEM is a system where all sciences are integrated." (T)

Theme 2: The need for STEM-integrated curriculum

During the interviews, a recurring theme was the importance of incorporating STEM into the curriculum. By including STEM applications, teachers would have sufficient time to implement them effectively. Furthermore, integrating STEM into the curriculum would create a need for assessments and the establishment of evaluation units within schools. This would encourage students and parents to take STEM seriously, leading to increased interest in the subject.

Here are some quotes that highlight the need for a STEM-integrated curriculum:

"As teachers, we have to follow a specific curriculum. So, the Ministry of National Education must prepare and plan a STEM curriculum to be integrated into teaching." (T)

"I don't think our education system is suitable for doing something directly related to STEM." (T)

"Here, there isn't enough time to encourage students to engage in STEM applications." (T)

"In order for unit and lesson plans to include STEM, a curriculum should be prepared for it." (A)

"The current curriculum is not suitable for STEM because it promotes an approach that encourages students to memorize." (T)

Theme 3: University exam anxiety is an obstacle for STEM

A significant challenge hindering STEM education is the university entrance examination, according to the participants. Students in the 11th and 12th grades are particularly preoccupied with this exam, leaving little time for other activities. Furthermore, there appears to be minimal pressure from parents and students to prioritize STEM education. Nevertheless, the participants recognize STEM as a fresh and exciting field—a sentiment that resonates with nearly all individuals surveyed.

Some quotes about this theme are provided below:

"After the 10th grade, our students prepare only for the university entrance exam. They are not interested in STEM because there are no STEM-related questions in the exam." (T)

"In order to be successful in STEM education, the examination system must be changed. Teachers and students should be motivated, and STEM-related goals should be made clear." (A)

"All students' lives revolve around exams. They just want to go to college." (T)

" Since the goal of the student is to be successful in the university entrance exam, the courses are taught as an information load." (T)

"The only thing parents want is for their children to have high grades and pass the university entrance exam." (A)

Theme 4: The need for budget and materials for STEM

Another recurring theme that emerged from the interviews was the necessity of having a budget specifically allocated for STEM education. The participants emphasized that while schools may have adequate physical facilities, the feasibility of STEM programs depends on the availability of necessary resources, which can only be ensured through a dedicated STEM budget.

Some quotes from participants about this theme are presented below:

"The school does not have the appropriate resources for STEM. There are no resources available for the students' academic level." (T)

"There are no materials we can use for mathematics at school." (T)

"We lack equipment. We can find funds not from the school but from some foundations when it is necessary for projects. Teachers sometimes pay out of pocket to buy resources." (A)

"There is no budget for direct STEM education. There is a budget for projects and experiments, but it is not sufficient for students' needs." (A)

Summary of the Scores

The readiness level of the regular public high school is categorized as "emerging" with a total score of 30 points, although it is closer to the lower boundary of the "early" category. In contrast, the vocational high school also falls under the "emerging" category with a total score of 44 points, positioning it near the threshold of the "integrated" level. Both the science high school and the private high school are classified at the "integrated" level, with scores that are nearly identical. The overall scores of each school are illustrated in Figure 4 and Figure 5.

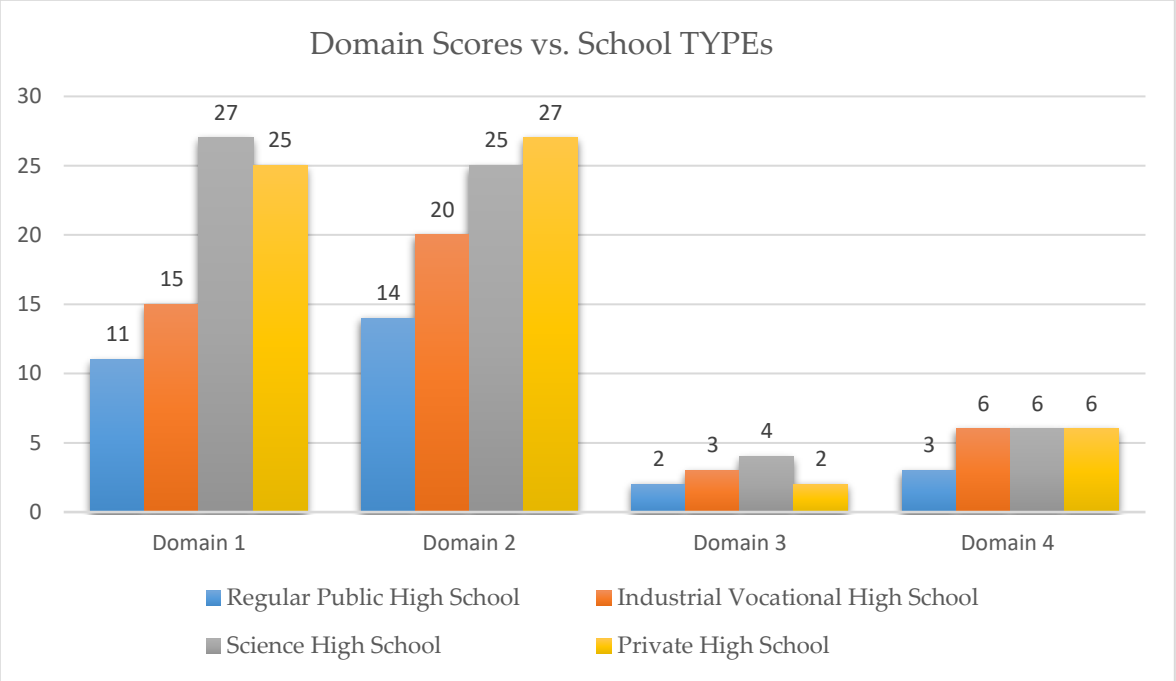


Figure 4. Domain Scores According to School Types

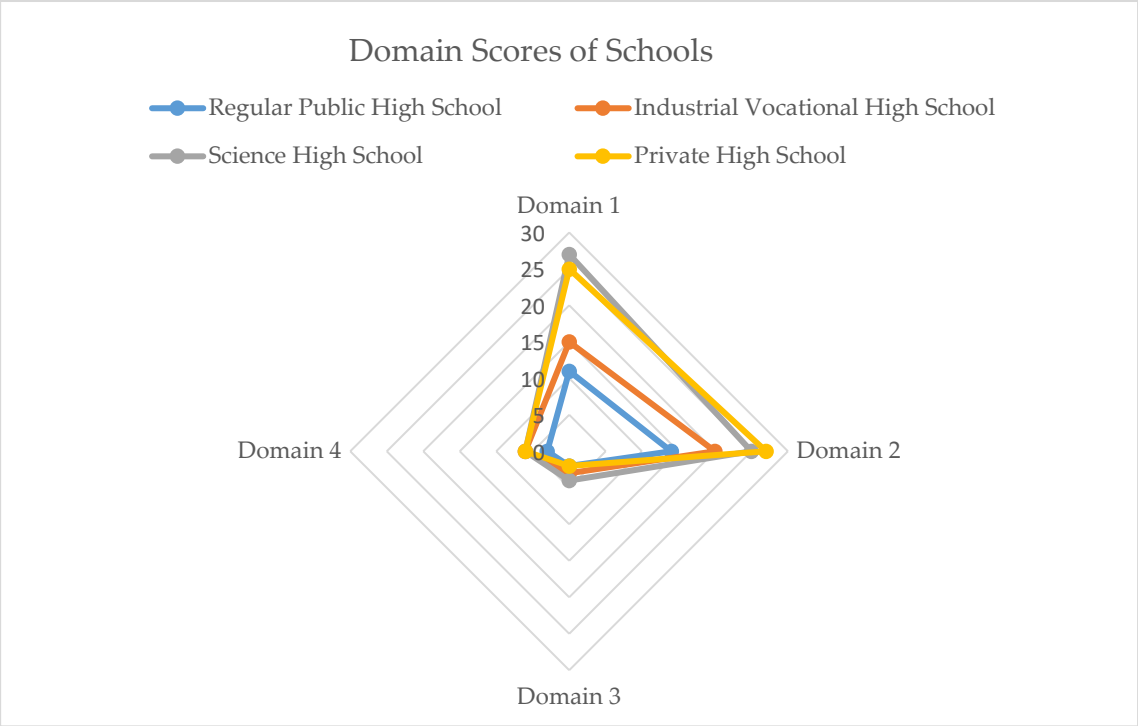


Figure 5. Summary of Results

Discussion

The present study examined STEM education readiness across four types of high schools in Ankara, revealing significant institutional disparities in cultural, instructional, and structural capacities. Although some schools demonstrated promising characteristics, none reached the fully integrated level, and all exhibited context-specific challenges reflecting both systemic barriers and localized constraints. These findings indicate that readiness for STEM education is not a uniform construct but a multidimensional condition shaped by institutional culture, teacher capacity, and policy context. Therefore, efforts to enhance STEM readiness must address both the broader structural inequalities among school types and the specific needs that arise within each local context.

At the regular public high school and vocational high school, although the physical infrastructure, including classrooms and laboratories, was found to be broadly adequate, the school's overall readiness for STEM integration was significantly hindered by material constraints, low student and parent interest, a rigid exam-focused academic culture, and a notable lack of teacher motivation. These findings support the work of Marginson et al. (2013), who emphasize that successful STEM implementation requires not only adequate facilities but also strong institutional leadership, pedagogical vision, and stakeholder engagement. Additionally, the exam-oriented nature of the Turkish education system reflects broader challenges faced by many high-performing yet high-pressure academic contexts. Le et al. (2021) assert that assessment-driven cultures diminish opportunities for creative, hands-on learning and create disincentives for both teachers and students to engage in STEM. This pattern is compounded when stakeholders do not perceive STEM as relevant or valuable, as observed by Baharin et al. (2018), who link low community awareness of STEM's long-term utility to decreased motivation and participation. These insights underscore the importance of addressing both the cultural mindset and instructional conditions alongside physical readiness.

The private high school, benefiting from the structure of the International Baccalaureate (IB) Programme, appeared better prepared for STEM implementation. Teachers demonstrated familiarity with interdisciplinary methods, and the curriculum already included STEM-aligned objectives. However, challenges such as the limited number of IB-enrolled students and faculty, as well as the persistent pressure of the university entrance examination, hindered full institutional alignment. Stella (2020) similarly observed that even in pedagogically progressive environments, high-stakes testing contexts create misalignments between curriculum goals and actual classroom priorities. Consequently, students and teachers may deprioritize inquiry-based learning in favor of exam preparation, thereby weakening the effectiveness of otherwise well-designed STEM curricula.

The science high school, although not formally affiliated with an international curriculum such as the IB Programme, achieved the highest STEM readiness score, indicating a strong institutional vision and structural capacity. However, exam pressure remained the most frequently cited obstacle among participants. This aligns with Ghaleb's (2024) analysis of high-performing secondary schools, which found that the dominance of university entrance examinations often narrows the curriculum, discourages interdisciplinary learning, and fosters heightened anxiety that is counterproductive to authentic STEM engagement. While teacher motivation was comparatively higher in this school, the lack of systemic support to balance exam preparation with exploratory learning emerged as a limiting factor.

Across all four schools, the average readiness score of 49 out of 100 places them at the "emerging" level according to the rubric used. Although the science and private high schools demonstrate promising institutional characteristics, systemic issues such as the national exam system, fragmented teacher training, underdeveloped partnerships, and minimal stakeholder engagement continue to hinder the full-scale integration of STEM. These findings align with the conclusions of Schweingruber et al. (2014) and English (2016), who argue that STEM readiness should be understood not merely as a function of resources but as a dynamic construct shaped by cultural beliefs, professional development systems, and alignment with broader educational policies.

These findings also align with the study's theoretical framework, which conceptualizes readiness as a multidimensional construct shaped by institutional culture, teacher capacity, and systemic policies. Prior research in Türkiye has highlighted similar challenges: Asiroglu and Akran (2018) reported that while teachers felt moderately competent in integrating STEM, they lacked readiness in areas such as engineering practices, technology-supported instruction, and innovative assessment. Likewise, Dönmez and Gülen (2023) emphasized that although science curricula have been frequently updated, teacher training programs have not kept pace, resulting in a persistent gap between curricular expectations and classroom realities. From a policy perspective, these results underscore the need for systemic investment in professional development, stronger coordination between MoNE and CoHE, and reforms aimed at reducing the dominance of high-stakes testing. A limitation of the present study is its focus on four schools in Ankara, which constrains the generalizability of the findings. Nevertheless, the multiple-case design provides valuable insights into the systemic and institutional factors shaping STEM readiness across diverse school types.

Conclusion

This study aimed to assess the readiness of four different types of high schools in Ankara to implement STEM education. The research was guided by the STEM Framework developed by the New York City Department of Education (NYCDOE, 2015), which emphasizes a holistic and interdisciplinary approach to science, technology, engineering, and mathematics. In line with global trends, Türkiye has become increasingly aware of the need to equip students with skills that foster innovation, problem-solving, and technological adaptability. Developing these competencies at the secondary school level is essential for preparing students to meet the demands of a rapidly changing world and for ensuring the country's long-term economic competitiveness.

Findings from the study revealed significant differences in readiness levels among the four school types. The regular public and vocational high schools demonstrated lower levels of preparedness, primarily due to inadequate resources, limited stakeholder awareness, low academic motivation, and the dominant influence of exam-oriented education. Teachers at these schools also reported limited engagement with STEM practices, often citing time constraints and the perceived difficulty of integrating STEM into the national curriculum. In contrast, the private and science high schools exhibited relatively higher levels of readiness, supported by more structured programs such as the International Baccalaureate (IB) Programme and stronger institutional visions. However, even in these schools, challenges such as exam pressure and the absence of a dedicated STEM curriculum remained significant barriers to full integration.

A key conclusion of the study is that readiness for STEM education cannot be achieved through infrastructure alone. It requires a combination of pedagogical, cultural, and systemic factors working together. Teacher capacity and training, clear curriculum alignment, administrative support, and collaboration among school stakeholders all play critical roles. Moreover, the widespread exam-oriented culture continues to hinder creativity, inquiry, and project-based learning opportunities that are central to effective STEM education. For meaningful progress to occur, there must be a shift in both mindset and practice across the education system.

To advance STEM integration, schools should no longer treat it as a supplementary program or an occasional enrichment activity. Instead, it should be viewed as a central component of the curriculum that supports interdisciplinary learning across subject areas. This transition requires concrete steps, including curriculum reform, investments in teacher professional development, improved resource allocation, and increased engagement with families and local communities. Efforts should also focus on raising awareness among all stakeholders about the long-term value of STEM for students' futures and the nation's development. A unified and collaborative approach will increase the likelihood of successful implementation.

In summary, although the current average readiness level of Ankara's high schools is classified as "emerging," there is strong potential for improvement if systemic barriers are addressed. The insights gained from this research provide a valuable foundation for future planning, policy-making, and school-level reform efforts. With coordinated action and a commitment to long-term development, high schools in Türkiye can make meaningful progress in STEM education. In doing so, they will not only better prepare students for 21st-century careers but also strengthen to the country's innovation capacity and global competitiveness.

Suggestions

The research aimed to assess the readiness of high schools for STEM education, with the potential to broaden future studies to include preschool, primary, secondary, and higher education institutions. The evaluation in this study was based on interviews with teachers and administrators, and future research could expand to include students and parents. Additionally, the findings from this study can serve as a valuable reference for future STEM reports, and the revised STEM framework may be applied in subsequent investigations.

This research has provided a comprehensive overview of high schools' readiness levels; however, future studies could examine additional factors such as teachers' age, gender, and families' socioeconomic status that may influence preparedness. Moreover, the findings suggest that high schools in Ankara are not yet adequately equipped for STEM education. Teachers' feedback highlights the need for curriculum developers to design an integrated STEM curriculum for high schools that includes specific objectives such as algorithm creation, machine construction, or model design.

To evaluate the effectiveness of STEM initiatives, higher authorities could conduct systematic assessments of schools implementing these programs. Additionally, to help schools obtain affordable and accessible instructional materials, such resources could be developed and made available for purchase. Finally, seminars and workshops could be organized to inform families about the importance of STEM education and to foster greater community interest and engagement.

References

Altunel, M. (2018). STEM eğitimi ve Türkiye: Fırsatlar ve riskler. *Seta Perspektif*, *207*(1), 1-7.

Anıl, D., Özkan, Ö. Y., & Demir, E. (2015). *PISA 2012 araştırması ulusal nihai rapor*. Millî Eğitim Bakanlığı. https://corumodm.meb.gov.tr/meb_iys_dosyalar/2018_11/03154501_pisa2012-ulusal-nihai-raporu.pdf

Arozaq, M., Aman, A., & Sunarhadi, M. (2017). Implementation of reading guide strategy in global climate change material for enhancement of student learning outcome. *International Journal of Active Learning*, *2*(2), 82-89.

Arslan, S. Y., & Arastaman, G. (2021). Dünyada STEM politikaları: Türkiye için çıkarımlar ve öneriler. *Nevşehir Hacı Bektaş Veli Üniversitesi SBE Dergisi*, *11*(2), 894-910. https://doi.org/10.30783/nevsosbilen.903115

Asiroglu, S., & Akran, S. K. (2018). The readiness level of teachers in science, technology, engineering and mathematics education. *Universal Journal of Educational Research*, *6*(11), 2461-2470. https://doi.org/10.13189/ujer.2018.061109

Australian Academy of Science. (2006). *Mathematics and statistics: Critical skills for Australia's future.* https://review.ms.unimelb.edu.au/FullReport2006.pdf

Baharin, N., Kamarudin, N., & Manaf, U. K. A. (2018). Integrating STEM education approach in enhancing higher order thinking skills. *International Journal of Academic Research in Business and Social Sciences*, *8*(7), 810-821. https://doi.org/10.6007/IJARBSS/v8-i7/4421

Balka, D. (2011). Standards of mathematical practice and STEM. In Math–science connector (pp. 6-8). School Science and Mathematics Association.

Baxter, P., & Jack, S. (2008). Qualitative case study methodology: Study design and implementation for novice researchers. *The qualitative report*, *13*(4), 544-559.

Bush, S. B., Edelen, D., Roberts, T., Maiorca, C., Ivy, J. T., Cook, K. L., Tripp, L. O., Burton, M., Alameh, S., Jackson. C., Mohr-Schroeder, M. J., Schroeder, D. C., McCurdy, R. P., & Cox Jr, R. (2024). Humanistic STE(A)M instruction through empathy: Leveraging design thinking to improve society. *Pedagogies: An International Journal*, *19*(1), 60-79. https://doi.org/10.1080/1554480X.2022.2147937

Breiner, J. M., Harkness, S. S., Johnson, C. C., & Koehler, C. M. (2012). What is STEM? A discussion about conceptions of STEM in education and partnerships. *School Science and Mathematics*, *112*(1), 3-11. https://doi.org/10.1111/j.1949-8594.2011.00109.x

Bybee, R. W. (2010). What is STEM education?. *Science*, *329*(5995), 996. https://doi.org/10.1126/science.1194998

Bybee, R. W. (2013). *The case for STEM education: Challenges and opportunities.* National Science Teaching Association Press.

Chachashvili-Bolotin, S., Milner-Bolotin, M., & Lissitsa, S. (2016). Examination of factors predicting secondary students' interest in tertiary STEM education. *International Journal of Science Education*, *38*(3), 366-390. https://doi.org/10.1080/09500693.2016.1143137

Crane, B. D. (2017). Teacher openness and prosocial motivation: Creating an environment where questions lead to engaged students. *Management Teaching Review*, *2*(1), 7-16. https://doi.org/10.1177/2379298116673838

Creswell, J. W., & Poth, C. N. (2016). *Qualitative inquiry and research design: Choosing among five approaches*. SAGE.

Cunningham, C. M., & Lachapelle, C. P. (2014). Designing engineering experiences to engage all students. In S. Purzer, J. Strobel, & M. Cardella (Eds.), *Engineering in pre-college settings: Synthesizing research, policy, and practices* (pp. 117-142). Purdue University Press. https://doi.org/10.2307/j.ctt6wq7bh.10

Çakır, R., & Ozan, C. E. (2018). FeTeMM etkinliklerinin 7. sınıf öğrencilerinin akademik başarıları, yansıtıcı düşünme becerileri ve motivasyonlarına etkisi. *Gazi Üniversitesi Gazi Eğitim Fakültesi Dergisi*, *38*(3), 1077-1100. https://doi.org/10.17152/gefad.346067

Çetin, F. (2016). The relationship between the classroom management competence of teachers and their attitudes towards the profession of teaching and job satisfaction. *Electronic Turkish Studies, 11*(3), 791-808. https://doi.org/10.7827/TurkishStudies.9285

Darling-Hammond, L. (2020). Accountability in teacher education. *Action in Teacher Education*, *42*(1), 60-71. https://doi.org/10.1080/01626620.2019.1704464

Denzin, N. K., & Lincoln, Y. S. (2011). *The SAGE handbook of qualitative research*. SAGE.

Dönmez, İ., & Gülen, S. (2023). Science curricula and science teachers training in Turkey: Past, present and future. *Necatibey Faculty of Education Electronic Journal of Science and Mathematics Education*, *17*(2), 929-962. https://doi.org/10.17522/balikesirnef.1371492

Dörnyei, Z. (2007). *Research methods in applied linguistics*. Oxford University Press.

Eisenhardt, K. M., & Graebner, M. E. (2007). Theory building from cases: Opportunities and challenges. *Academy of Management Journal*, *50*(1), 25-32. https://doi.org/10.5465/amj.2007.24160888

English, L. D. (2016). STEM education K-12: Perspectives on integration. *International Journal of STEM Education*, *3*(1), 3. https://doi.org/10.1186/s40594-016-0036-1

Farina, C., Weinberg, P., Commitante, A., & Curtis-Bey, L. (2016). *NYC STEM education framework.* New York City Department of Education.

Fensham, P. J. (2008). *Science education policy-making: Eleven emerging issues*. UNESCO.

Ghaleb, B. D. S. (2024). Effect of exam-focused and teacher-centered education systems on students' cognitive and psychological competencies. *International Journal of Multidisciplinary Approach Research and Science*, *2*(2), 611-631. https://doi.org/10.59653/ijmars.v2i02.648

Ginott, H. (1976). *Teacher and child: A book for parents and teachers*. Macmillan.

Gonzalez, H. B., & Kuenzi, J. J. (2012, August). *Science, technology, engineering, and mathematics (STEM) education: A primer* (CRS Report No. R42470). Congressional Research Service, Library of Congress.

Hall, G. E., & Hord, S. M. (2011). *Implementing change: Patterns, principles, and potholes* (3rd ed.). Pearson.

Harman, G., & Çelikler, D. (2012). Eğitimde hazır bulunuşluğun önemi üzerine bir derleme çalışması. *Eğitim ve Öğretim Araştırmaları Dergisi, 1*(3), 147-156.

Hernandez, P. R., Bodin, R., Elliott, J. W., Ibrahim, B., Rambo-Hernandez, K. E., Chen, T. W., & de Miranda, M. A. (2014). Connecting the STEM dots: Measuring the effect of an integrated engineering design intervention. *International Journal of Technology and Design Education*, *24*(1), 107-120. https://doi.org/10.1007/s10798-013-9241-0

Hersey, P., & Blanchard, K. H. (1988). *Management of organizational behavior: Utilizing human resources* (5th ed.). Prentice Hall.

Hsieh, H. F., & Shannon, S. E. (2005). Three approaches to qualitative content analysis. *Qualitative Health Research*, *15*(9), 1277-1288. https://doi.org/10.1177/1049732305276687

Kolb, S. M. (2012). Grounded theory and the constant comparative method: Valid research strategies for educators. *Journal of Emerging Trends in Educational Research and Policy Studies*, *3*(1), 83-86.

Kondracki, N. L., Wellman, N. S., & Amundson, D. R. (2002). Content analysis: Review of methods and their applications in nutrition education. *Journal of Nutrition Education and Behavior*, *34*(4), 224-230. https://doi.org/10.1016/S1499-4046(06)60097-3

Le, H. C., Nguyen, V. H., & Nguyen, T. L. (2023). Integrated STEM approaches and associated outcomes of K-12 student learning: A systematic review. *Education Sciences*, *13*(3), 297. https://doi.org/10.3390/educsci13030297

Le, L. T. B., Tran, T. T., & Tran, N. H. (2021). Challenges to STEM education in Vietnamese high school contexts. *Heliyon*, *7*(12), e08649. https://doi.org/10.1016/j.heliyon.2021.e08649

Lynch, D., & Smith, R. (2016). Readiness for school reform. *International Journal of Innovation, Creativity and Change*, *2*(3), 1-12.

Marginson, S., Tytler, R., Freeman, B., & Roberts, K. (2013). *STEM: Country comparisons*. Australian Council of Learned Academies.

Margot, K. C., & Kettler, T. (2019). Teachers' perception of STEM integration and education: A systematic literature review. *International Journal of STEM Education*, *6*(1), 1-16. https://doi.org/10.1186/s40594-018-0151-2

McDonald, C. (2016). STEM education: A review of the contribution of the disciplines of science, technology, engineering and mathematics. *Science Education International*, *27*(4), 530-569.

Merriam-Webster. (1999). *Merriam-Webster's collegiate dictionary* (10th ed.).

Miles, M. B., & Huberman, M. (1994). *Qualitative data analysis: An expanded sourcebook.* Thousand Oaks.

Ministry of National Education. (2016a). *PISA 2015 Türkiye ulusal raporu*. Millî Eğitim Bakanlığı Ölçme, Değerlendirme ve Sınav Hizmetleri Genel Müdürlüğü. https://odsgm.meb.gov.tr/meb_iys_dosyalar/2023_10/09155512_PISA_2015.pdf

Ministry of National Education. (2016b). *STEM education report.* https://yegitek.meb.gov.tr/STEM_Education_Report.pdf

Ministry of National Education. (2024). *PISA 2022 Türkiye ulusal raporu*. Millî Eğitim Bakanlığı Ölçme, Değerlendirme ve Sınav Hizmetleri Genel Müdürlüğü. https://pisa.meb.gov.tr/meb_iys_dosyalar/2024_03/21120745_26152640_pisa2022_rapor.pdf

Moore, T. J., Tank, K. M., Glancy, A. W., & Kersten, J. A. (2015). NGSS and the landscape of engineering in K-12 state science standards. *Journal of Research in Science Teaching*, *52*(3), 296- https://doi.org/10.1002/tea.21199

Morrison, J. (2006). *Attributes of STEM education: The student, the school, the classroom*. Teaching Institute for Excellence in STEM. https://www.leadingpbl.org/f/Jans%20pdf%20Attributes_of_STEM_Education-1.pdf

Mueller, J. (2008). Assessing skill development. *Library Media Connection*, *27*(3), 18-20.

National Research Council. (2012). *A framework for K-12 science education: Practices, crosscutting concepts, and core ideas.* The National Academies Press. https://doi.org/10.17226/13165

NYC Department of Education. (2015). *NYC STEM education framework.* https://www.weteachnyc.org/resources/resource/stem-framework/

Oral, I., & McGivney, E. (2013). *Türkiye'de matematik ve fen bilimleri alanlarında öğrenci performansı ve başarının belirleyicileri TIMSS 2011 analizi.* Eğitim Reformu Girişimi. https://egitimreformugirisimi.org/yayinlar/timss-2011-analizi/

Ozturk, N., & Roehrig, G. H. (2025). Effects of an integrated STEM unit designed around socioscientific issues on middle school students' socioscientific reasoning. *International Journal of Science and Mathematics Education, 23*(5), 1493-1518. https://doi.org/10.1007/s10763-024-10517-8

Radloff, J., & Guzey, S. (2016). Investigating preservice STEM teacher conceptions of STEM education. *Journal of Science Education and Technology, 25*(5), 759-774. https://doi.org/10.1007/s10956-016-9633-5

Schunk, D. H. (2012). Behaviorism. In P. Smith (Ed.), *Learning theories: An educational perspective* (pp. 63-76). Pearson Education.

Schweingruber, H., Pearson, G., & Honey, M. (2014). *STEM integration in K-12 education: Status, prospects, and an agenda for research.* National Academies Press.

Shim, S. W. (2024). Cultivating empathy in STEM: A conceptual framework for multicultural empathic design in STEM education. *Journal of STEM Teacher Education, 59*(1), 2. https://doi.org/10.61403/2158-6594.1531

Sinclair, C. (2008). Initial and changing student teacher motivation and commitment to teaching. *Asia-Pacific Journal of Teacher Education, 36*(2), 79-104. https://doi.org/10.1080/13598660801971658

Solheim, K., Ertesvåg, S. K., & Dalhaug Berg, G. (2018). How teachers can improve their classroom interaction with students: New findings from teachers themselves. *Journal of Educational Change, 19*(4), 511-538. https://doi.org/10.1007/s10833-018-9333-4

Stake, R. E. (1995). *The art of case study research.* SAGE.

Stella, M. (2020*).* Forma mentis networks reconstruct how Italian high schoolers and international STEM experts perceive teachers, students, scientists, and school. *Education Sciences, 10*(1), 17. https://doi.org/10.3390/educsci10010017

Şahin, A., Ayar, M. C., & Adıgüzel, T. (2014). STEM-related after-school program activities and associated outcomes on student learning. *Educational Sciences: Theory and Practice, 14*(1), 309-322. https://doi.org/10.12738/estp.2014.1.1876

Şanlı, M., & Özerbaş, D. H. S. (2021). STEM etkinliklerinin öğrencilerin STEM alanlarına yönelik tutumuna ve fene yönelik motivasyonlarına etkisi. *Manisa Celal Bayar Üniversitesi Sosyal Bilimler Dergisi, 19*(3), 139-154. https://doi.org/10.18026/cbayarsos.889816

TUSIAD. (2017). *STEM alanında Türkiye'nin insan kaynağı ihtiyacı 2023.* https://tusiad.org/tr/yayinlar/raporlar/item/9735-2023-e-dog-ru-tu-rkiye-de-stem-gereksinimi

Vermunt, J. D. (2014). Teacher learning and professional development. In S. Krolak-Schwerdt, S. Glock, & M. Böhmer (Eds.), *Teachers' professional development: Assessment, training, and learning* (pp. 79-95). Sense Publishers. https://doi.org/10.1007/978-94-6209-536-6_6

Education and Science

Original Article

Vol 51 (2026) No 226 405-420

Comparison of guided and unguided internet-based cognitive behavioral therapy for test anxiety

Cumhur Avcil [1], Oğuzhan Herdi [2]

Abstract

This study aims to compare the effectiveness of guided and unguided internet-based cognitive behavioral therapies (iCBT) in reducing test anxiety. A total of 64 participants were included in the study, with 32 in the guided group and 32 in the unguided group. Test anxiety levels before and after the intervention were measured using the Test Anxiety Inventory (TAI), and adherence was assessed based on the completion percentage of the iCBT program. Data were analyzed using repeated measures analysis of variance. A significant main effect was found for the time factor ($F(1, 62) = 108.046$, $p < .001$), but the time × group interaction was not significant ($F(1, 62) = 1.153$, $p = .287$). Holm-adjusted post hoc analyses revealed significant and large effect size reductions in test anxiety scores in both groups (unguided group: $d = 1.562$; guided group: $d = 1.922$). No significant difference was observed between the groups' final test scores ($p = .832$, $d = 0.153$), and the effect size difference was also not significant ($z = 0.63$, $p = .528$). Furthermore, the percentage of guided group members using iCBT was significantly higher than that of the unguided group ($U = 336.000$, $p = .017$), but no significant relationship was found between the percentage of use and the change in test anxiety ($r = .127$, $p = .317$). The findings indicate that guided intervention has the potential to increase user adherence, while unguided iCBT may also be an effective and independently applicable intervention option. In conclusion, both intervention formats offer effective, accessible, and flexible psychological support options for reducing test anxiety.

Keywords

Test anxiety

Internet-based cognitive behavioral therapy

Cognitive behavioral therapy

Article Info

Received: 11.06.2023

Accepted: 11.03.2025

Published Online: 04.30.2026

DOI: 10.15390/ES.2026.2463

Introduction

It is a natural reaction for individuals to experience a certain level of stress and anxiety in any situation where they are required to perform in their daily lives. Exams that assess academic achievement are one such situation where these types of emotional responses are commonly observed. Under intense expectations and performance pressure, individuals' anxiety levels may increase, preventing them from effectively utilizing their current abilities and fully demonstrating their potential

[1] Antalya Science University, Faculty of Economics, Administrative and Social Sciences, Department of Psychology, Antalya, Türkiye, cumhur.avcil@antalya.edu.tr

[2] Antalya Science University, Faculty of Economics, Administrative and Social Sciences, Department of Psychology, Antalya, Türkiye, oguzhan.herdi@antalya.edu.tr

(Doherty & Wenderoth, 2017). In this context, test anxiety is considered a type of performance anxiety. Test anxiety shares common characteristics with other types of performance anxiety, such as sports and public speaking anxiety (American Psychiatric Association [APA], 2021).

Test anxiety is a type of anxiety that causes a decline in academic performance because of emotional, behavioral, and physiological responses arising from the individual's cognitive processes during the test preparation process or during the test itself. Evaluations of test anxiety in the literature are consistent with our definition and support the view that this condition can negatively affect an individual's academic success (Dusek, 1980; Spielberger & Vagg, 1995; Zeidner, 2020).

Studies on the prevalence of test anxiety report that a significant proportion of students experience intense anxiety about exams. While the prevalence of test anxiety in Western societies ranges between 25-40% (Thomas et al., 2018; Zeidner, 2020), studies conducted in different cultures show that the prevalence of test anxiety among students is around 25-30% (Yıldırım, 2008). Research conducted in Turkey, however, reports higher levels of test anxiety. Yıldırım (2007) stated that 42% of students preparing for exams in Turkey experienced high test anxiety, while Kavakci et al. (2014) reported this rate to be 48%. Furthermore, it is stated that approximately 18% of students with high test anxiety have their academic achievement negatively affected by high anxiety levels (Adana & Kaya, 2005).

Interventions implemented to improve test anxiety are generally based on cognitive behavioral therapy (CBT) (Ergene, 2003). CBT is an evidence-based therapy method that combines cognitive and behavioral techniques and has been shown to be effective in reducing students' test anxiety (Larson et al., 2010; Neuderth et al., 2009; Putwain & von der Embse, 2021; Ramirez & Beilock, 2011; von der Embse et al., 2013). However, some studies suggest that cognitive interventions may be effective in reducing test anxiety even without behavioral techniques (Ulusoy et al., 2016). Nevertheless, Ergene's (2003) comprehensive meta-analysis shows that combining skill-focused interventions with cognitive and behavioural techniques is more effective in alleviating test anxiety.

Internet-based cognitive behavioral therapy (iCBT) is an intervention method that provides clients with the structured content of face-to-face CBT sessions, making it easy to use (Andersson, 2016). iCBT teaches cognitive behavioral techniques to its users through interactive software that can be used on devices such as mobile phones, tablets, or computers, and encourages users to apply these techniques regularly (Carlbring et al., 2018). The effectiveness of iCBT has been researched on various psychological disorders, and successful results have been obtained in many studies. Studies on diagnoses such as depression (Clarke et al., 2005; Cuijpers et al., 2010; Melling & Houguet-Pincham, 2011), anxiety (Dryman et al., 2017; Jakobsen et al., 2017; Nordgren et al., 2014), and social anxiety (Kampmann et al., 2016; Klein et al., 2010) have shown that iCBT is effective in treating these disorders.

This study examines two types of iCBT programs (guided and unguided iCBT). In guided iCBT programs, the therapist interacts with the client through face-to-face, live video sessions, phone calls, or messaging to guide the process (Titov et al., 2010). The therapist monitors progress, while the software takes on the tasks of teaching new CBT techniques and providing practical exercises (Andersson & Titov, 2014). On the other hand, unguided iCBT programs offer a structure that does not require the therapist's direct participation and can be completed by users without professional support (Baumeister et al., 2014).

A recent meta-analysis including randomized controlled trials has shown that guided and unguided iCBT are equally effective in treating anxiety and related disorders (Oey et al., 2023). The results indicate that guided and unguided interventions lead to similar clinical outcomes, but guided interventions may be more effective in the short term. However, follow-up studies have observed that the advantages of guided interventions diminish over time. Furthermore, it was determined that the number of sessions in guided interventions increased the effect size, while reminders and intervention duration had no significant effect in unguided interventions (Oey et al., 2023).

Another meta-analysis comparing guided and unguided iCBT indicated that severity of distress is also an important determinant (Karyotaki et al., 2021). This study indicated that guided iCBT was more effective than unguided iCBT, but this difference was particularly pronounced in individuals with moderate to severe depression. In contrast, unguided iCBT was found to have similar levels of effect in individuals with mild depressive symptoms. Furthermore, both guided and unguided iCBT were found to be more effective than control treatments in the short and long term, but the superiority of guided iCBT decreased significantly at 6- or 12-month follow-ups.

Looking at studies applying iCBT for test anxiety, it is seen that research in this area is quite limited compared to other psychological disorders. Orbach et al. (2007) conducted the first randomized controlled trial involving iCBT interventions in the modern sense. The results of the study show that the iCBT program can significantly reduce test anxiety. Compared to the control group, a significant decrease in test anxiety levels was observed in the intervention group. In the literature, a guided iCBT study involving limited therapist intervention (one written message per week) found that students in the experimental group showed a significant decrease in test anxiety. Furthermore, it has been reported that students participating in the intervention experienced a significant decrease in their overall psychological distress and anxiety levels (Warnecke et al., 2020). These findings suggest that iCBT may be an effective intervention method for improving test anxiety among university students. One of the limited number of studies conducted in Turkey examined the effect of an unguided iCBT program on students preparing for high school and university exams (Avcil & Herdi, 2023). The study found that students who used the iCBT program experienced a significant decrease in their test anxiety levels, and their state and trait anxiety levels also decreased significantly compared to the control group.

The limited number of studies in the literature on iCBT interventions for test anxiety and the recommendations in existing studies for further research necessitate the expansion of scientific knowledge in this area. Accordingly, this study aims to evaluate the effect of an iCBT program developed for test anxiety and to compare guided and unguided applications. It is expected that a guided iCBT study conducted specifically with a Turkish sample will reveal whether it offers an alternative approach to traditional CBT and whether iCBT is an effective tool for reducing test anxiety.

Within this scope, two main hypotheses were tested in the study: (H1) The reduction in test anxiety among students using the guided iCBT program will be greater than that among students using the unguided iCBT program; (H2) Both types of intervention are predicted to provide a significant reduction compared to the initial levels of test anxiety. To evaluate these hypotheses, changes in test anxiety among students who used the iCBT program through face-to-face sessions guided by a therapist were compared with changes in test anxiety scores among students who used the iCBT program on their own without guidance.

Method

This study examined the effectiveness of guided and unguided implementation of an iCBT program developed for test anxiety. A 2x2 experimental design was used in the study. The first factor represents the independent treatment groups (guided and unguided iCBT), while the second factor represents repeated measurements of the dependent variable (pre-test, and post-test). The dependent variable of the study was test anxiety level, while the independent variable was the guided or unguided iCBT intervention program applied to the participants.

Participants

Participants in the guided iCBT group were individuals who sought help at a private psychiatric clinic in Antalya for test anxiety complaints, had no psychiatric diagnosis other than test anxiety complaints, and did not require medication. Participants in the unguided iCBT group registered on a website (psikademi.com) designed to offer iCBT programs free of charge and used the test anxiety program (No More Test Anxiety) on their own. Participants were asked the questions, "Have you ever received a psychiatric diagnosis?" and "Are you taking any psychiatric medication?". All participants answered these questions and stated that they did not have any psychiatric diagnoses and were not

taking any medication. Based on these statements, they were included in the study. The inclusion criteria for the study were; being a student, scoring above 40 on the Test Anxiety Inventory (TAI), volunteering to participate in the study, not having any psychiatric diagnosis, not using psychiatric medication, and having used at least 60% of the iCBT program (14 days). Informed consent forms were obtained from all participants over the age of 18 and from both the participants themselves and at least one parent for participants under the age of 18.

Participants were assigned to groups based on their preferred condition (face-to-face, online). A total of 64 participants were included in the study: 32 in the guided iCBT group and 32 in the unguided iCBT group. A total of 98 individuals registered for the program before being included in the unguided iCBT group, but 66 of these participants were excluded from the analysis because they did not meet the criterion of completing at least 60% of the program. On the other hand, no exclusions were made in the guided iCBT group because all participants who applied for the study met the inclusion criteria. The average age of the guided group was 17.40, and the group consisted of 24 girls and 8 boys. The average age of the unguided group was 16.90, and the group consisted of 24 girls and 8 boys. Of the participants in the guided iCBT group, 26 were preparing for the Higher Education Institutions Exam (YKS), 5 for the High School Transition Exam (LGS), and 1 for the Vertical Transition Exam (DGS). Of the participants in the unguided iCBT group, 20 were preparing for the YKS, 10 for the LGS, and 1 for the DGS (Table 1). The equalization of participant numbers and gender distribution in the groups occurred not as a planned intervention but as a natural consequence of the data collection process.

Table 1. Participants' sociodemographic data

		Guided iCBT Group		**Unguided iCBT Group**	
		M	**sd**	**M**	**sd**
Age		17.406	2.448	16.906	2.787
		n	**%**	**n**	**%**
Gender	Female	24	75	24	75
	Male	8	25	8	25
School	Middle School	5	15.62	10	31.25
	High School	26	81.25	20	62.50
	University	1	3.12	2	6.25
Prepared Exam*	LGS	5	15.62	10	31.25
	YKS	26	81.25	20	62.50
	DGS	1	3.12	1	3.12
Total		32		32	

*LGS: High School Transition Exam, YKS: Higher Education Institutions Exam, DGS: Vertical Transfer Exam

Data Collection Tools

Demographic Information Form: The demographic information form consists of questions designed to collect basic information about participants, including their age, gender, city of residence, school, the exam they are preparing for, and whether they have received any psychological/psychiatric treatment.

Test Anxiety Inventory: Test anxiety was measured in this study using the Test Anxiety Inventory. The inventory was developed by Spielberger in 1980 and adapted into Turkish, with validity and reliability studies conducted (Öner, 1990). The scale consists of 20 Likert-type items, and total scores range from 20 to 80. High scores indicate high test anxiety, while low scores indicate low test anxiety. The Cronbach's alpha coefficient for the internal consistency of the scale was found to be 0.87 in the Turkish sample (Öner, 1990). In this study, the Cronbach's alpha coefficient for the TAI was calculated as .94.

Treatment Procedure

The iCBT procedure (No More Test Anxiety) developed for test anxiety is primarily a 21-day web-based program consisting of online videos, online tests, and exercises aimed at changing thoughts. The program includes psychoeducation, cognitive restructuring, cognitive and behavioral exercises aimed at creating study habits and increasing and maintaining motivation, exercises aimed at changing attention and thinking focus, breathing and relaxation exercises, and exposure content related to test anxiety. The first 14 days of the program introduce participants to new methods and skills, while the last 7 days are dedicated to reinforcing these skills. Therefore, a minimum completion threshold of 60% (14 days) has been set as the criterion for inclusion in the study. Users can access the program via smartphone, tablet, or computer and complete the daily content in approximately 15 minutes. Appendix 1 details the program's daily content and flow.

Prior to the intervention, participants in the unguided iCBT group did not undergo any face-to-face psychological or psychiatric assessment. This choice was based on the assumption that, in an online and voluntary intervention structure, referring participants to a healthcare facility could create access difficulties and lead to sample loss. Participants were informed that they would use the program on their own and were encouraged to participate in the process without receiving any guidance in person, online, via email, or text message. Participants' progress throughout the program was monitored solely through the website's administrator panel. However, no additional assessment was made at any stage of the process regarding whether participants received other psychological support.

Participants in the guided iCBT group underwent an individual assessment interview before starting the process, and situations in which they experienced test anxiety were assessed through face-to-face interviews. After the initial measurements were completed, participants were included in the intervention for test anxiety using the iCBT program. Throughout this process, the guidance service was provided solely by a clinical psychologist, and all sessions with participants were conducted one-on-one. The first author of the study took on the guidance role. The first author holds a master's degree in clinical psychology and a doctorate in psychology and also has training and clinical practice experience in cognitive behavioral therapy. Although involved in the program's development phase, standard procedures were adhered to throughout the counseling process to minimize the risk of bias. Sessions were conducted using structured content, and measurements were collected via an independent, automated online system. The role of the counselor was not to directly provide psychotherapy but to inform participants about the techniques and methods included in the iCBT program and to support the effective use of the program. During the sessions, explanations were provided about the cognitive behavioral therapy techniques used to cope with test anxiety, but the applications were not provided directly during the sessions; rather, they were provided through exercises that participants completed via the iCBT program. Between sessions, participants were given homework assignments that included cognitive restructuring, relaxation techniques, motivation enhancement exercises, and exposure exercises included in the program. Participants' completion rates and progress were monitored through the website's administrator panel. The number of face-to-face sessions for participants in the guided iCBT group varied according to individual needs. The average number of sessions was 6, with a minimum of 3 and a maximum of 18 sessions conducted. Session duration and frequency were determined based on the participant's commitment to the process and individual needs. The average intervention duration for participants was 13 weeks, with the shortest duration being 3 weeks and the longest being 32 weeks.

Participants in the guided group completed an average of 89.46% of the iCBT program (range: 60%-100%), and half of the participants (n=16) completed the program 100%. In contrast, participants in the unguided iCBT group participated in the program without any therapist or guide support. Participants completed the program on their own over 30 days, with a fixed 30-day interval between the initial and final measurements. It was determined that the unguided group completed an average of 83% of the iCBT program (range: 60%-100%), with half of the participants (n=16) completing more than 85% of the program and 3 participants completing 100%.

The final measurements for all participants in both groups were taken before they took the exam they had prepared for, and no additional measurements were taken after they took the exam. Therefore, variables such as exam results and exam performance were not evaluated in the study.

During the study, care was taken to apply the same iCBT intervention to participants in both groups. The intervention consisted of standard components such as cognitive restructuring, relaxation techniques, motivation enhancement exercises, and exposure exercises, and was presented in the same format for both groups. Participants in the guided iCBT group received additional guidance from a clinical psychologist while using the program, but no direct psychotherapy was administered during this process. The role of the guide was limited to informing participants about the techniques included in the iCBT program, increasing their adherence to the process, and guiding the use of the program. Accordingly, although the same intervention content was offered to both groups, there were some structural differences between the groups apart from the guidance support. The guided group consisted of participants identified through clinical referral and undergoing a preliminary assessment process, while the unguided group was determined based on online application and included in the process without any face-to-face assessment. This adds an additional dimension to the study in terms of comparing the applicability of intervention types in real-life conditions.

Findings

The study included a total of 64 participants, 32 in the guided iCBT group and 32 in the unguided iCBT group. Before proceeding to the analyses of Test Anxiety Inventory (TAI) scores, the assumptions of normal distribution and variance homogeneity were evaluated. The Shapiro-Wilk normality test results were $W = 0.974$, $p = .186$ for the first measurement of the TAI and $W = 0.969$, $p = .108$ for the final measurement, indicating that the assumption of normal distribution was met for both measurements. Levene's test results were $F(1, 62) = 0.019$, $p = .890$ for the first TAI measurement and $F(1, 62) = 3.600$, $p = .062$ for the final measurement, confirming that the variances were homogeneous. The iCBT usage percentage data were found to be non-normally distributed, and the differences between groups were analyzed using the non-parametric Mann-Whitney U test. To evaluate the effect of the intervention on test anxiety, a two-factor analysis of variance with repeated measures was applied, with time (pre-test - post-test) as the first factor and group (guided - unguided) as the second factor.

The Mann-Whitney U test was applied to assess whether the iCBT usage percentage differed between the guided and unguided iCBT groups (Table 2). The analysis results show that the usage percentage of the guided iCBT group is significantly higher than that of the unguided group ($U = 336.000$, $p = .017$). The mean usage percentage was calculated as 83.44 ($SD = 12.09$) for the unguided group and 89.47 ($SD = 13.48$) for the guided group. Figure 1 shows each participant's individual usage percentages (dots), the median and quartile values for each group (box plot), and the distribution densities (curves) together. The guided group generally exhibits a higher and more stable usage rate. The effect size of this difference was found to be $r = .344$ using the rank-biserial correlation coefficient and was assessed as corresponding to a medium effect size. However, since no significant relationship was observed between the usage percentage and the change in test anxiety ($r = .127$, $p = .317$), this variable was not included as a control variable in subsequent analyses.

Table 2. Comparison of iCBT usage percentage between groups

Measure	Group	N	M	Standard Deviation	Standard Error	Coefficient of Variation	U	p
iCBT	Unguided	32	83.438	12.091	2.137	0.145	336.000	.017
Usage Percentage	Guided	32	89.469	13.484	2.384	0.151		

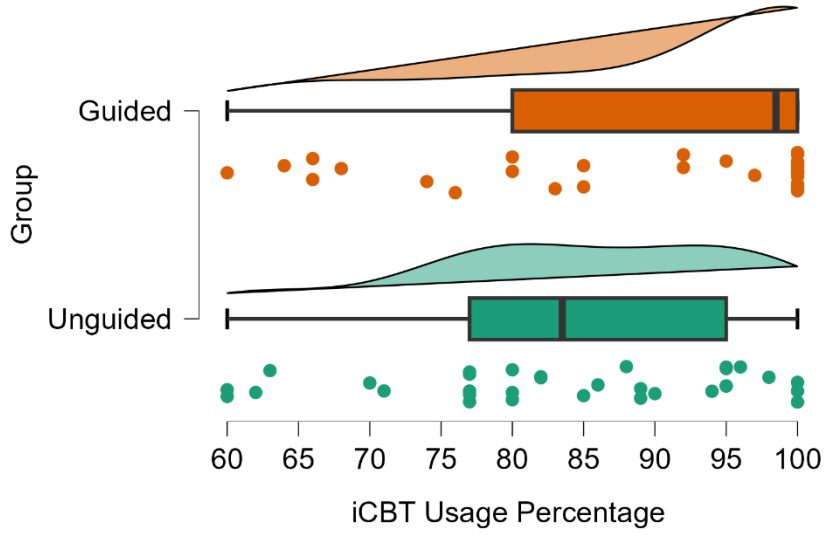


Figure 1. Distribution of iCBT usage percentage among groups

Repeated measures analysis of variance results (Table 3) indicates that the intervention program (pre-test - post-test) had a significant effect on test anxiety levels ($F(1, 62) = 108.046$, $p < .001$). This finding indicates that test anxiety decreased significantly in all participants after the intervention. However, the time × group interaction is not statistically significant ($F(1, 62) = 1.153$, $p = .287$), suggesting that the decrease in anxiety levels after the intervention occurred similarly in the guided and unguided groups. Furthermore, there was no significant difference in overall test anxiety levels between groups ($F(1, 62) = 0.021$, $p = .884$).

Table 3. Repeated measures ANOVA results

Measurements	Group	N	M	Standard Deviation	Standard Error	Coefficient of Variation
TAI Pre-Treatment	Unguided	32	54.281	9.247	1.635	0.170
	Guided	32	56.188	9.386	1.659	0.167
TAI Post-Treatment	Unguided	32	39.906	10.858	1.919	0.272
	Guided	32	38.500	6.886	1.217	0.179
Residuals		**Sum of Squares**	**df**	**Mean Square**	**F**	**p**
Time		8224.031	1	8224.031	108.046	**<.001**
Time × Group		87.781	1	87.781	1.153	0.287
Residuals		4719.188	62	77.116		
Groups		2.000	1	2.000	0.021	0.884
Residuals		5786.875	62	93.337		

*TAI: Test Anxiety Inventory

Holm-adjusted post hoc comparisons (Table 4) showed no significant difference between the initial measurement scores of the guided and unguided groups ($p_{holm} = .832$, $d = -0.207$, 95% CI = [-0.899, 0.484]). On the other hand, a significant decrease was observed between the initial and final measurement scores of both groups. In the unguided group, this difference was $M_{difference} = 14.375$, $p_{holm} < .001$, $d = 1.562$, 95% CI = [0.811, 2.312], while in the guided group, $M_{difference} = 17.688$, $p_{holm} < .001$, $d = 1.922$, 95% CI = [1.123, 2.721]. When the final measurement scores of the groups were compared, no significant difference was observed ($p_{holm} = .832$, $d = 0.153$, 95% CI = [-0.521, 0.827]). A z test comparing the effect sizes (Cohen's d) indicated that the difference was not statistically significant ($z = 0.63$, $p = .528$). These results indicate that both interventions were effective in reducing test anxiety,

but no significant difference emerged between the groups after treatment. The distribution of TAI scores for guided and unguided iCBT groups is shown in Figure 2.

Table 4. Post hoc analysis results

Measurements		Mean Difference	Standard Error	t	Cohen's d	95% CI for Cohen's d		pholm
						Low	High	
Unguided Initial	Guided Initial	-1.906	2.329	-0.818	-0.207	-0.899	0.484	0.832
	Unguided Final	14.375	2.181	6.591	1.562	0.811	2.312	**<.001**
	Guided Final	15.781	2.301	6.858	1.714	0.914	2.515	**<.001**
Guided Initial	Unguided Final	16.281	2.301	7.075	1.769	0.962	2.576	**<.001**
	Guided Final	17.688	2.181	8.109	1.922	1.123	2.721	**<.001**
Unguided Final	Guided Final	1.406	2.273	0.619	0.153	-0.521	0.827	0.832

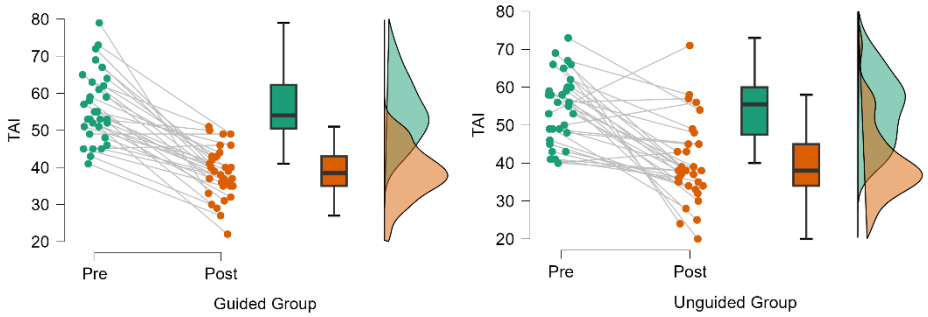


Figure 2. Distribution of TAI scores for guided and unguided iCBT groups

Finally, the relationships between the total number of sessions spent with the therapist, the weekly duration of the intervention, and the percentage of iCBT platform usage for participants in the guided iCBT group and the change in test anxiety (difference in TAI scores) were examined using Pearson correlation analysis. The analyses revealed that these variables did not show a significant relationship with the change in test anxiety ($p > .05$ for all correlations). Similarly, no significant correlation was found between the percentage of use and the change in test anxiety in the unguided iCBT group ($p > .05$).

Discussion

This study aimed to compare the effectiveness of guided and unguided iCBT in reducing test anxiety. Repeated measures analysis of variance results showed that the time factor (pre-treatment - post-treatment) had a significant effect on test anxiety scores, but the time × group interaction was not significant. This finding indicates that both intervention forms are effective in reducing test anxiety, but that guidance support does not provide a statistically significant additional contribution to anxiety levels after the intervention. Comparing the findings with other studies in literature, evaluating the original contributions of the study, discussing possible limitations, and presenting suggestions for future research are addressed in this section.

This study demonstrated that both guided and unguided iCBT are effective in reducing test anxiety. These findings are consistent with previous studies showing that CBT-based interventions are successful in reducing test anxiety (Ergene, 2003; Putwain & von der Embse, 2021; Ramirez & Beilock,

2011). It is known that CBT components such as cognitive restructuring, relaxation techniques, and exposure are effective on cognitive distortions and physiological responses, which are the core components of test anxiety (von der Embse et al., 2013). An iCBT study on test anxiety (Orbach et al., 2007) showed that this form of intervention could significantly reduce test anxiety compared to the control group. Consistent with these preliminary findings, the current study once again confirms that internet-based interventions can be an effective method for test anxiety. When examining the effects of guided and unguided iCBT on reducing test anxiety, it was found that both intervention methods significantly reduced test anxiety. According to repeated measures ANOVA results, the effect of iCBT was statistically significant, but the interaction between iCBT and group was not significant. Holm-adjusted post hoc comparisons indicated marked pre- to post-treatment reductions in test anxiety in both groups, yet no significant between-group difference at post-test. Furthermore, the z-test conducted to assess the significance of the difference between these two effect sizes was not statistically significant. This finding indicates that assumptions regarding the contribution of guidance support to the effect of the iCBT intervention are not statistically supported, at least in the sample of this study.

Previous studies have examined the effectiveness of guided and unguided iCBT in different ways. Warnecke and colleagues' (Warnecke et al., 2020) study showed that guided iCBT is effective in reducing test anxiety and general psychological distress. However, in the study, the guided iCBT intervention was administered over six weeks with regular therapist support. In this study, the duration of guided iCBT varied based on individual differences, and therapist support was limited to providing guidance and direction rather than direct psychotherapy. These differences may explain the contrasting results between the two studies. Correlation analyses revealed no significant relationship between the number of sessions, duration, and percentage of use and changes in test anxiety. Furthermore, it was observed that unguided iCBT also significantly reduced test anxiety. This finding is consistent with previous studies showing that unguided iCBT can be effective on its own (Baumeister et al., 2014; Titov et al., 2010). Indeed, a recent study also reported that unguided iCBT significantly reduced test anxiety and that there was a significant negative correlation between program adherence (usage rate) and final test anxiety scores (Avcil & Herdi, 2023). These findings suggest that unguided iCBT can be effective even without therapist support and that user adherence may play a decisive role in clinical outcomes.

In this study, the guided iCBT group completed the program at a higher rate. The higher completion/adherence observed in guided formats aligns with prior literature (Andersson & Titov, 2014; Karyotaki et al., 2021) and may be attributable to process-level components such as regular feedback, motivational check-ins, and problem-solving support. These adherence differences may reflect process-level support effects rather than superiority in symptom reduction. Accordingly, even in the absence of a statistically significant between-group post-test difference, guided delivery may enhance implementation fidelity and sustainability in routine clinical practice. Furthermore, it has been reported that dropout rates are generally high in unguided iCBT applications and that this may be related to individual variables such as age, education level, relationship status, and initial symptom severity (Karyotaki et al., 2015). The fact that our study mainly included young participants at the high school level may help explain the high dropout rate observed in the unguided group. However, since participants' initial anxiety levels were already high due to inclusion criteria, it was not possible in this study to assess the effect of symptom severity on program dropout. Nevertheless, this study found no significant relationship between usage percentage and change in test anxiety. This finding suggests that user adherence alone may not be sufficient to determine therapeutic gains.

This study contributes significantly to the literature by comparing the effectiveness of guided and unguided internet-based cognitive behavioral therapies (iCBT) in reducing test anxiety. Using a comparative experimental design and repeated measures ANOVA, effects related to both intervention type (guided/unguided) and time (pre-treatment/post-treatment) were systematically evaluated. Furthermore, adherence levels to the intervention process were measured through the percentage of iCBT usage, and a significant effect of guided intervention on adherence was demonstrated. However, the lack of a significant relationship between usage rate and change in test anxiety suggests that the role

of adherence in therapeutic outcomes may be more complex. Studies directly comparing guided and unguided iCBT, analyzing both program usage rates and symptom change, are quite limited. This experimental study conducted with a Turkish sample is also important in that it contributes to the limited number of studies examining the effect of guided iCBT in particular. In this respect, the current study provides empirical data on different forms of iCBT in the field of test anxiety and guides practitioners in developing flexible options in line with individual differences and intervention preferences. From the practitioners' perspective, when choosing between online intervention formats, it should be considered that unguided iCBT can yield effective results with fewer resources.

However, the study has some limitations. First, this study examined only post-intervention effects. To draw conclusions about long-term outcomes, the effects of guided and unguided iCBT should be measured in the following months. In this context, not only symptom reduction but also broader outcomes such as the persistence of acquired cognitive and behavioral skills, relapse rates, and functionality should be evaluated. It is known that CBT-based interventions provide sustainable long-term benefits by equipping individuals with coping strategies. However, it is not yet clear to what extent these strategies are durable when delivered online, whether guided or unguided. Second, since neither group in this study received traditional face-to-face therapy, the findings are based on a direct comparison of online therapy formats. Guided iCBT does not fully reflect the traditional therapeutic process but instead involves limited guidance and directional support. Therefore, the nature of the relationship and the forms of support involved when the guided group had contact with the guide for a longer period have not been evaluated in detail. Future studies should comprehensively investigate the relationship between the quality and content of time spent with the guide and symptom change. Third, individual psychological characteristics that could affect adherence to the intervention and therapeutic gains were not measured in this study. This is a limitation that should be considered when interpreting the findings. Fourth, the counseling service was provided by one of the researchers. Although the researcher's expertise in clinical psychology and experience in applying CBT increased the objectivity of this process, this situation can be considered a potential source of bias. Finally, the sample of this study is limited to students preparing for national high-stakes central examinations. This limits the generalizability of the findings to other age groups, academic levels, or cultural contexts. Similar studies with populations with different characteristics will provide broader and more reliable results regarding the effects of iCBT on test anxiety.

This study demonstrated that both guided and unguided iCBT are effective in reducing test anxiety. Although the findings revealed that guided intervention had a higher effect size, there was no statistically significant difference between the two methods. This indicates that unguided iCBT can also be considered a powerful alternative for reducing test anxiety. Furthermore, the higher completion rate among participants in the guided iCBT group highlights the positive effect of therapist support on user adherence. However, considering that participants in the unguided iCBT group also achieved significant gains, it is thought that unguided iCBT could be promoted as a self-help-based intervention.

The findings of this study provide important contributions to interventions aimed at managing test anxiety; however, further research is needed to evaluate the effects of guided and unguided iCBT in a broader context. The findings obtained in this study indicate that online applications are used more intensively and regularly by users when conducted with guidance. Therefore, it would be beneficial to consider the guidance element in the design of programs to be developed in the future. Developing strategies to increase user engagement and integrating new technological support to sustain the effect of unguided iCBT should be one of the focal points of future studies. In conclusion, while this study demonstrates that iCBT is an effective tool for reducing test anxiety, it emphasizes the need for a more in-depth examination of how individual differences are shaped depending on the type of intervention.

Conclusion

This study has demonstrated that guided and unguided internet-based cognitive behavioral therapies (iCBT) are effective intervention methods in reducing test anxiety. Significant reductions in test anxiety levels were observed in both groups; however, no statistically significant difference was found between the groups in terms of final test scores. Although the effect size was higher in the guided iCBT group, this difference was not statistically significant.

Analyses of the program usage level indicated that the guided intervention form may increase user adherence; however, no significant relationship was found between the usage percentage and the decrease in test anxiety. These findings suggest that unguided iCBT may also be effective and may offer a cost-effective intervention option, especially in conditions where resources are limited. Although guided iCBT has positive effects on user engagement, the role of engagement in therapeutic outcomes appears to be more complex.

Future studies are recommended to focus on more comprehensive and follow-up research designs aimed at monitoring the long-term sustainability of guided and unguided iCBT applications and their varying effectiveness according to individual differences.

References

Adana, F., & Kaya, N. (2005). Lise öğrencilerinin sınav kaygısı düzeyi üzerine sınav kaygısı ile başa çıkma eğitiminin etkisi. *Kriz Dergisi, 13*(2), 35-42.

American Psychiatric Association. (2021, April 20). *What are anxiety disorders?.* https://www.psychiatry.org/patients-families/anxiety-disorders/what-are-anxiety-disorders

Andersson, G. (2016). Internet-delivered psychological treatments. *Annual Review of Clinical Psychology, 12*(1), 157-179. https://doi.org/10.1146/annurev-clinpsy-021815-093006

Andersson, G., & Titov, N. (2014). Advantages and limitations of Internet-based interventions for common mental disorders. *World Psychiatry, 13*(1), 4-11. https://doi.org/10.1002/wps.20083

Avcil, C., & Herdi, O. (2023). An unguided computerized cognitive behavioral therapy program for test anxiety. *Journal of Cognitive-Behavioral Psychotherapy and Research, 12*(3), 231-241. https://doi.org/10.5455/JCBPR.135830

Baumeister, H., Reichler, L., Munzinger, M., & Lin, J. (2014). The impact of guidance on Internet-based mental health interventions—A systematic review. *Internet Interventions, 1*(4), 205-215. https://doi.org/10.1016/j.invent.2014.08.003

Carlbring, P., Andersson, G., Cuijpers, P., Riper, H., & Hedman-Lagerlöf, E. (2018). Internet-based vs. face-to-face cognitive behavior therapy for psychiatric and somatic disorders: An updated systematic review and meta-analysis. *Cognitive Behaviour Therapy, 47*(1), 1-18. https://doi.org/10.1080/16506073.2017.1401115

Clarke, G., Eubanks, D., Reid, E., Kelleher, C., DeBar, L., Lynch, F., & Nunley, S. (2005). Overcoming depression on the Internet (ODIN) (2): A randomized trial of a self-help depression skills program. *Journal of Medical Internet Research, 7*(2), e16. https://doi.org/10.2196/jmir.7.2.e16

Cuijpers, P., Donker, T., van Straten, A., Li, J., & Andersson, G. (2010). Is guided self-help as effective as face-to-face psychotherapy for depression and anxiety disorders? A systematic review and meta-analysis of comparative outcome studies. *Psychological Medicine, 40*(12), 1943-1957. https://doi.org/10.1017/S0033291710000772

Doherty, J. H., & Wenderoth, M. P. (2017). Implementing an expressive writing intervention for test anxiety in a large college course. *Journal of Microbiology & Biology Education, 18*(2), 18.2.39. https://doi.org/10.1128/jmbe.v18i2.1307

Dryman, M. T., McTeague, L. M., Olino, T. M., & Heimberg, R. G. (2017). Evaluation of an open-access CBT-based Internet program for social anxiety: Patterns of use, retention, and outcomes. *Journal of Consulting and Clinical Psychology, 85*(10), 988-999. https://doi.org/10.1037/ccp0000232

Dusek, J. B. (1980). The development of test anxiety in children. In I. G. Sarason (Ed.), *Test anxiety: Theory, research and applications* (pp. 87-110). Lawrence Erlbaum.

Ergene, T. (2003). Effective interventions on test anxiety reduction: A meta-analysis. *School Psychology International, 24*(3), 313-328. https://doi.org/10.1177/01430343030243004

Jakobsen, H., Andersson, G., Havik, O. E., & Nordgreen, T. (2017). Guided Internet-based cognitive behavioral therapy for mild and moderate depression: A benchmarking study. *Internet Interventions,* (7), 1-8. https://doi.org/10.1016/j.invent.2016.11.002

Kampmann, I. L., Emmelkamp, P. M. G., & Morina, N. (2016). Meta-analysis of technology-assisted interventions for social anxiety disorder. *Journal of Anxiety Disorders,* (42), 71-84. https://doi.org/10.1016/j.janxdis.2016.06.007

Karyotaki, E., Efthimiou, O., Miguel, C., genannt Bermpohl, F. M., Furukawa, T. A., Cuijpers, P., & Forsell, Y. (2021). Internet-based cognitive behavioral therapy for depression: A systematic review and individual patient data network meta-analysis. *JAMA Psychiatry, 78*(4), 361-371. https://doi.org/10.1001/jamapsychiatry.2020.4364

Karyotaki, E., Kleiboer, A., Smit, F., Turner, D. T., Pastor, A. M., Andersson, G., Berger, T., Botella, C., Breton, J. M., Carlbring, P., Christensen, H., de Graaf, E., Griffiths, K., Donker, T., Farrer, L., Huibers, M. J. H., Lenndin, J., Mackinnon, A., Meyer, B., … Cuijpers, P. (2015). Predictors of treatment dropout in self-guided web-based interventions for depression: An individual patient data meta-analysis. *Psychological Medicine, 45*(13), 2717-2726. https://doi.org/10.1017/S0033291715000665

Kavakci, O., Semiz, M., Kartal, A., Dikici, A., & Kugu, N. (2014). Test anxiety prevalence and related variables in students who are going to take the university entrance examination. *Dusunen Adam The Journal of Psychiatry and Neurological Sciences, 27(*4), 301-307. https://doi.org/10.5350/DAJPN2014270403

Klein, B., Mitchell, J., Abbott, J., Shandley, K., Austin, D., Gilson, K., & Redman, T. (2010). A therapist-assisted cognitive behavior therapy Internet intervention for posttraumatic stress disorder: Pre-, post-, and 3-month follow-up results from an open trial. *Journal of Anxiety Disorders, 24(*6), 635-644. https://doi.org/10.1016/j.janxdis.2010.04.005

Larson, H., Ramahi, M., Conn, S., Estes, L., & Ghibellini, A. (2010). Reducing test anxiety among third grade students through the implementation of relaxation techniques. *Journal of School Counseling,* 8(19), 1-19.

Melling, B., & Houguet-Pincham, T. (2011). Online peer support for individuals with depression: A summary of current research and future considerations. *Psychiatric Rehabilitation Journal, 34*(3), 252-254. https://doi.org/10.2975/34.3.2011.252.254

Neuderth, S., Jabs, B., & Schmidtke, A. (2009). Strategies for reducing test anxiety and optimizing exam preparation in German university students: A prevention-oriented pilot project of the University of Würzburg. *Journal of Neural Transmission, 116*(6), 785-790. https://doi.org/10.1007/s00702-008-0123-7

Nordgren, L. B., Hedman, E., Etienne, J., Bodin, J., Kadowaki, Å., Eriksson, S., & Carlbring, P. (2014). Effectiveness and cost-effectiveness of individually tailored Internet-delivered cognitive behavior therapy for anxiety disorders in a primary care population: A randomized controlled trial. *Behaviour Research and Therapy,* (59), 1-11. https://doi.org/10.1016/j.brat.2014.05.007

Oey, L. T., McDonald, S., McGrath, L., Dear, B. F., & Wootton, B. M. (2023). Guided versus self-guided Internet-delivered cognitive behavioral therapy for diagnosed anxiety and related disorders: A preliminary meta-analysis. *Cognitive Behaviour Therapy, 52*(6), 654-671. https://doi.org/10.1080/16506073.2023.2250073

Orbach, G., Lindsay, S., & Grey, S. (2007). A randomized placebo-controlled trial of a self-help Internet-based intervention for test anxiety. *Behaviour Research and Therapy, 45*(3), 483-496. https://doi.org/10.1016/j.brat.2006.04.002

Öner, N. (1990). *Sınav kaygısı envanteri el kitabı.* Yüksek Öğrenimde Rehberliği Tanıtma ve Rehber Yetiştirme Vakfı Yayını.

Putwain, D. W., & von der Embse, N. P. (2021). Cognitive-behavioral intervention for test anxiety in adolescent students: Do benefits extend to school-related wellbeing and clinical anxiety?. *Anxiety, Stress, and Coping, 34*(1), 22-36. https://doi.org/10.1080/10615806.2020.1800656

Ramirez, G., & Beilock, S. L. (2011). Writing about testing worries boosts exam performance in the classroom. *Science, 331*(6014), 211-213. https://doi.org/10.1126/science.1199427

Spielberger, C. D., & Vagg, P. R. (1995). *Test anxiety: Theory, assessment, and treatment.* Taylor & Francis.

Thomas, C. L., Cassady, J. C., & Finch, W. H. (2018). Identifying severity standards on the Cognitive Test Anxiety Scale: Cut score determination using latent class and cluster analysis. *Journal of Psychoeducational Assessment, 36*(5), 492-508. https://doi.org/10.1177/0734282916686004

Titov, N., Andrews, G., Davies, M., McIntyre, K., Robinson, E., & Solley, K. (2010). Internet treatment for depression: A randomized controlled trial comparing clinician vs. technician assistance. *PLoS One, 5*(6), e10939. https://doi.org/10.1371/journal.pone.0010939

Ulusoy, S., Yavuz, K. F., Esen, F. B., Umut, G., & Karatepe, H. T. (2016). Cognitive group therapy for test anxiety. *Journal of Cognitive Behavioral Psychotherapy and Research,* (1), 28-37. https://doi.org/10.5455/JCBPR.212133

von der Embse, N., Barterian, J., & Segool, N. (2013). Test anxiety interventions for children and adolescents: A systematic review of treatment studies from 2000-2010. *Psychology in the Schools, 50*(1), 57-71. https://doi.org/10.1002/pits.21660

Warnecke, I., Haffinger, L. M., Konradi, H., Reiss, N., Luka-Krausgrill, U., & Rohrmann, S. (2020). Design of a guided Internet-delivered counseling intervention for test anxiety. *The European Journal of Counselling Psychology, 8*(1), 163-176. https://doi.org/10.5964/ejcop.v8i1.205

Yıldırım, İ. (2007). Depression, test anxiety, and social support among Turkish students preparing for the university entrance examination. *Eurasian Journal of Educational Research, 7(*29), 171-188.

Yıldırım, İ. (2008). Family variables influencing test anxiety of students preparing for the university entrance examination. *Eurasian Journal of Educational Research*, *8*(31), 171-186.

Zeidner, M. (2020). Test anxiety. In B. J. Carducci, C. S. Nave, A. Di Fabio, D. H. Saklofske, & C. Stough (Eds.), *The Wiley encyclopedia of personality and individual differences* (pp. 445-449). John Wiley & Sons. https://doi.org/10.1002/9781119547174.ch253

Appendix 1. Content of the iCBT Program for Test Anxiety (No More Test Anxiety)

The "No More Test Anxiety" program is a 21-day program that includes cognitive restructuring, changing and organizing study behaviors, thought and behavior exercises to increase and maintain motivation, attention and thought focus exercises, breathing and relaxation exercises, and imaginary exposure exercises. Students log into the program daily via mobile phone, tablet, or computer and complete the assigned exercises for that day in approximately 15 minutes. The program's web address is www.psikademi.com.

The detailed structure of the program is as follows:

Day 1: In addition to providing information about the use of the iCBT program, the results of the test anxiety inventory are explained. Then, what test anxiety is, and its causes are explained through psycho-educational videos based on the cognitive model.

Day 2: Cognitive training videos explaining the relationship between anxiety and thoughts based on the cognitive model and the anxiety avoidance relationship are provided. Detailed information is then given about automatic thoughts and cognitive distortions, and participants are taught how to identify and record automatic thoughts.

Day 3: Begins with an interactive quiz based on identifying cognitive distortions through common thoughts seen in test anxiety. After this exercise, users are informed about generating alternatives to automatic thoughts and how to counter them, and forms are provided for them to complete this exercise. The final activity of the day involves psychoeducation about breathing exercises and their implementation.

Day 4: Begins with an interactive quiz that starts to generate alternatives to common automatic thoughts in test anxiety. Then, access is provided to thought forms that the person can use on their own, explaining how to change these thoughts and how to achieve realistic and functional outcomes. The day ends with psychoeducation on ways to cope with avoidance and a video on homework.

Day 5: Begins with an interactive exercise requiring the generation of realistic and functional thoughts. This is followed by a video explaining effective study methods for exams. The final exercise of the day concludes with an explanation and practice of progressive relaxation exercises.

Day 6: Begins with an exercise involving written work aimed at generating realistic thoughts. This is followed by psychoeducation on study planning. The final activity of the day is the explanation and practice of relaxation and guided imagery exercises.

Day 7: Includes refilling the exam anxiety inventory and explaining what observed or unobserved changes mean. Finally, what has been done up to this stage is summarized, and subsequent homework assignments are explained.

Day 8: Begins with psychoeducation on motivation and ways to overcome motivation problems. This is followed by an explanation and application of mindfulness-based sensory focus exercises.

Day 9: Begins with psychoeducation on coping with not enjoying work. This is followed by an attention-focusing exercise using visual stimuli that distracts the person from physical stimuli.

Day 10: Begins with psychoeducation explaining solutions for indecision between doing other things or working. This is followed by a written daily planning exercise. The day's final activity concludes with exercises to shift thought focus.

Day 11: There is psychoeducation explaining daily behaviors such as nutrition, physical exercise, and sleep that are effective in reducing anxiety. Then, psychoeducation is provided on how to use the work done up to this stage when anxiety occurs.

Day 12: Common mistakes students make during exams and their solutions, along with exam tactics, are discussed. Afterwards, there is a video on what to do to cope with anxiety the night before the exam.

Day 13: The day begins with a video explaining what to do to cope with anxiety on the morning of the exam. Afterwards, it explains how to use the methods studied so far during the exam.

Day 14: The most important activity of the day is explaining what exposure through imagination is and performing an exposure. The scenario created for this exposure has a flow like the central high school and university entrance exams in Turkey. This day also includes filling out the exam anxiety inventory and explaining the results.

Day 15: It begins with an informative video explaining what will happen in the next phase. This is followed by an interactive quiz aimed at identifying cognitive distortions and a breathing exercise application.

Day 16: Includes an interactive quiz based on generating alternatives to automatic thoughts and a progressive relaxation exercise.

Day 17: Includes an interactive quiz aimed at achieving realistic and functional outcomes and a guided imagery exercise for relaxation.

Day 18: Includes a written exercise based on producing realistic and functional outcomes and a mindfulness exercise based on focusing on the senses.

Day 19: Includes exercises to shift the person's attention and thought focus away from thoughts and physical stimuli.

Day 20: Includes an exposure exercise involving an imagination exercise with an exam scenario.

Day 21: Includes videos on completing the exam anxiety scale one last time and interpreting the results. Finally, there is a psycho-educational video containing information on what students can do in cases where the program is insufficient in reducing exam anxiety or in cases of recurrence.

Education and Science

Original Article

Vol 51 (2026) No 226 421-439

The relationship between teachers' critical thinking disposition and teachers' autonomy support: a moderated mediation model of teacher autonomy and self-efficacy

İbrahim Çolak [1], Ahmet Şakir Yazıcı [2]

Abstract

Teachers' autonomy support has attracted increased scholarly attention due to its positive effects on students and educational outcomes. This study examined a moderated mediation model to explore the underlying mechanisms linking critical thinking disposition and teachers' autonomy support. In this study, we proposed teacher autonomy as a mediator and self-efficacy as a moderator of the relationship between critical thinking disposition and teachers' autonomy support. Data collected from 417 teachers in Türkiye were analyzed utilizing bootstrapping tests. The results revealed both significant direct and indirect effects of critical thinking disposition on teachers' autonomy support through teacher autonomy. Self-efficacy significantly moderated the relationship between critical thinking disposition and teacher autonomy. In addition, self-efficacy significantly moderated the strength of the mediated relationship between critical thinking disposition and autonomy support. When teachers experienced higher self-efficacy, the positive indirect effects of critical thinking disposition on teachers' autonomy support were found to be stronger than on counterparts who experienced low self-efficacy. Our discussion emphasizes the benefits of comprehending the functions of teacher autonomy and self-efficacy in contributing to the effectiveness of critical thinking disposition on teachers' autonomy support. Our findings provide practical recommendations for enhancing teacher autonomy and self-efficacy.

Keywords

Critical thinking
Autonomy support
Teacher autonomy
Self-efficacy
Moderated mediation
Teacher

Article Info

Received: 11.04.2024
Accepted: 11.17.2025
Published Online: 04.30.2026

DOI: 10.15390/ES.2026.2473

Introduction

Contemporary educational environments should be designed in a way that empowers students to act on from an internal locus of causality and have a sense of choice over their learning activities. Such an environment can be achieved by autonomy supportive practices implemented by teachers. In the most general sense, autonomy refers to the internal endorsement of actions, characterized by high volition and the flexibility in self-regulating one's behavior (Reeve & Jang, 2006). In the literature

[1] Trabzon University, Fatih Faculty of Education, Department of Educational Sciences, Trabzon, Türkiye, ibrahimcolak@trabzon.edu.tr
[2] Muğla Provincial Directorate of National Education, Muğla, Türkiye, ahmetsakiryazici@gmail.com

autonomy occupies a central place in Self-Determination Theory (SDT), as proposed by Ryan and Deci (2000, 2017). According to this theory, autonomy, competence, and relatedness are among the factors that facilitate or hinder individual's flourishing. Therefore, autonomy is regarded as one of the basic psychological needs that promote motivation, growth, and well-being. In this view, autonomy is not equivalent to independence or individualism. Instead, it corresponds to individuals self-regulating their behaviors in alignment with their authentic values and interests (Ryan & Deci, 2017).

The development of autonomy from the early years of life is affected by the approaches of teacher and parents, whose behavior can be either controlling or autonomy supportive (Deci & Ryan, 2000). Teachers who are autonomy-supportive engage and foster students' psychological needs, preferences, and interests. SDT has also been widely investigated in terms of educational setting and the research has emphasized that students receiving more autonomy support from their teachers demonstrated more autonomous motivation, competence, and performance (Black & Deci, 2000). Some research showed links between autonomy supportive teaching and students' satisfaction, well-being, and higher grades (Sheldon & Krieger, 2007). In addition, the satisfaction of autonomy has been found to enhance the intrinsic motivation and thereby performance of the students (Guay et al., 2001). As can be seen, teachers' autonomy support is recognized as a pivotal element in ensuring the success of learners. International organizations such as Organization for Economic Co-operation and Development (OECD) and European Union also underline the importance of autonomy supportive environments in education. For example, in order to promote lifelong learning, European Union (2019) emphasizes creating learning environments that foster the development of key competencies, such as entrepreneurship and learning to learn, which can be promoted by supporting students' autonomy. In addition, OECD (2018) underlines autonomy as one of the skills that students should acquire to shape and thrive in their future world. Thus, exploring the underlying factors of teachers' autonomy support can provide a better understanding of the concept, its connections with other variables, and the mechanisms through which it can be cultivated.

Recent literature has extensively investigated teachers' autonomy support, including its antecedents (Basri, 2023; Reeve & Cheon, 2016) and its outcomes (Jang et al., 2010; Shen et al., 2009). There are also studies that have examined the relationships between teachers' autonomy support and self-efficacy beliefs (Akçil & Oğuz, 2015), teacher autonomy (Yazıcı, 2016), and educational beliefs (Oğuz et al., 2014). Although limited research has tested the link between critical thinking disposition and teachers' autonomy support (Bolaños-Medina & Núñez, 2022; Koçoğlu & Kanadlı, 2019; Sevari & Farzadi, 2022), they were conducted with samples drawn from students and did not employ teachers' autonomy support as dependent variable. In addition, to the authors' knowledge, no prior research examined teacher autonomy as a mediator and self-efficacy as a moderator in this relationship. In order to address this gap in the literature, this study aimed to examine the mediating role of teacher autonomy and moderating role of self-efficacy in the relationship between teachers' critical thinking disposition and their autonomy support. The current study adds to the literature in certain aspects. First, the study developed a unique moderated mediation model of teacher autonomy support affected by critical thinking disposition, including teacher autonomy as a mediating variable and self-efficacy as a moderating variable (see Figure 1). By considering the mediating effect of teacher autonomy and moderating effect of self-efficacy, we strived to enhance the understanding of the underlying mechanisms and limiting factors that influence the connection between critical thinking disposition and teacher autonomy support. Second, this study extends the scope of the literature on teacher autonomy by explaining the conditions under which teachers' critical thinking disposition may influence their autonomy differently.

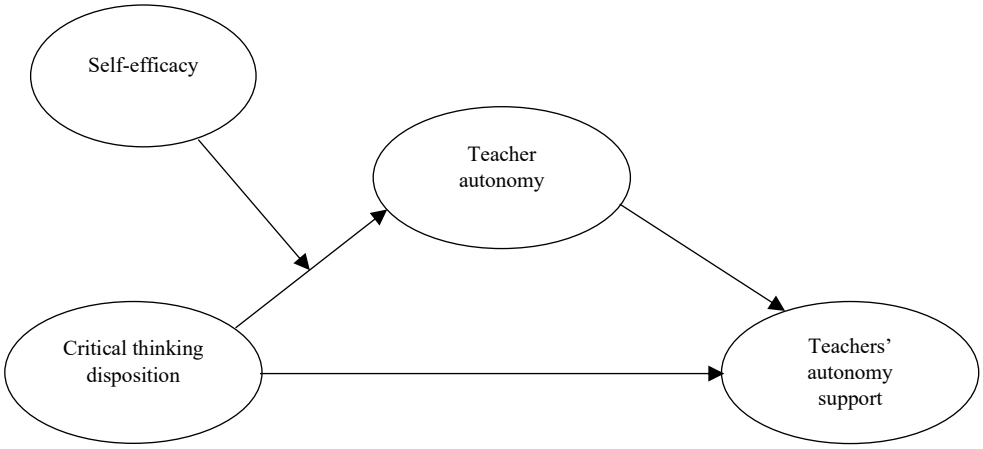


Figure 1. Conceptual model illustrating moderated mediation

Conceptual Framework

Teachers' Autonomy Support

Teachers' autonomy support is defined as teachers displaying behaviors in their professional work that build and actuate their students' sense of autonomy and internal motivation resources (Huéscar Hernández et al., 2020). Autonomy support involves teachers supporting learners to make appropriate choices and to minimize the pressure that their students may experience in fulfilling their duties, and supporting the personal initiative of learners rather than fostering a controlling environment in which external rewards and punishment are the essence (Deci & Ryan, 1991). Autonomy support also requires creating an environment in which students are not pressured to behave in a desired way, but are instead encouraged to be themselves, which in turn helps to develop their internal motivation (Ryan & Deci, 2004). Supporting learner autonomy is considered to have multifaceted benefits, such as enhancing the quality of education (Reeve & Jang, 2006) by helping students develop a positive attitude towards learning and enhancing their capacity to learn (Little, 1995). Teachers' autonomy support also promotes students' satisfaction, well-being, and academic achievement (Sheldon & Krieger, 2007; Tunçeli et al., 2022). A literature review conducted by Yang et al. (2022) demonstrated that teachers' autonomy support served as a foundation for student engagement and motivation. A study by Alkın-Şahin et al. (2015) found a positive and moderate correlation between classroom teachers' autonomy support and their critical thinking supportive behaviors. In addition, Núñez et al. (2015) concluded that teachers' autonomy support is a significant predictor of students' autonomy. Thus, the literature emphasizes positive outcomes for students as well as for the overall teaching environment when autonomy is supported.

In the academic setting, an autonomy-supportive environment is characterized by the following: Teachers understand and accept students' decisions, offer alternative solutions to problems and the opportunity to choose between different activities, and avoid controlling feedback or the use of explanatory logic (Núñez et al., 2015; Su & Reeve, 2011). On the other hand, the establishment of the agenda, providing solutions and answers, asking controlling questions, praising as a contingent reward, and criticizing the students are among the controlling behaviors exhibited by teachers. For a better learning environment, it is essential for educators to embrace an autonomy-supportive approach in their teaching (Reeve et al., 2004). However, this style requires a shift from the teacher's role as a conveyer of knowledge to that of a facilitator of learning and a manager of learning resources. Teachers need to determine areas in which they can develop learner autonomy. In other words, teachers must decide to what degree their students can decide their own learning goals, select their own learning materials, and contribute to the methods used by students to self-evaluate their own learning (Little, 1995). In order to achieve this, it is considered that teachers should be instructional leaders who follow the changes and developments in their profession, remain free from prejudice when making decisions, are willing to solve problems with their students in the learning environment, and to offer options to their students as and where possible and appropriate.

Critical Thinking Disposition and Teachers' Autonomy Support

Critical thinking is described as the process of questioning one's own thinking as well as that of others, involving a set of cognitive abilities and tendencies supporting decision making and problem solving in different situations (Çolak et al., 2022; Kılıç & Şen, 2014). According to another definition, critical thinking is regarded as an advanced kind of reflective thought that entails an increased awareness of individuals' perceptions, emotions, and actions (Phan, 2008). It can be characterized as one's internal drive in problem solving and decision making (Zhang & Lambert, 2008), and includes both emotional and cognitive components (Salsali et al., 2013). It refers to self-regulated judgment in making evaluation, inference, interpretation, and analysis in the cognitive system (Facione, 1990). In other words, while ordinary thinking is based upon guesswork, preferences, and assumptions, critical thinking is based on prediction, evaluation, and logical inference (Çolak et al., 2022). A critical thinker is an individual characterized by curiosity, extensive knowledge, a reliance upon logic, open-mindedness, flexibility, honesty in addressing personal biases, caution in making judgments, a willingness to rethink, diligence in seeking relevant information, as well as an aptitude for problem clarity and the ability to manage complex situations (Salsali et al., 2013). In particular, teachers with strong critical thinking skills are expected to go beyond their traditional teaching roles by supporting their students' independent learning and encouraging their autonomous actions.

Teachers' autonomy supportive behaviors encompass a set of instructional practices such as identifying students' needs, permitting students' interests and preferences to direct classroom activities, providing a meaningful rationale for rules and expectations, being responsive to students' questions, and offering performance-based praise (Reeve & Jang, 2006). This shows that autonomy supportive behaviors require teachers to have attributes such as flexibility, openness, the ability to create opportunities, innovativeness, cognitive skills, the capacity to stimulate motivation, and sensitivity to students' needs. Therefore, to foster student autonomy, teachers need to employ higher-order thinking abilities, which are also inherent in critical thinking. Correspondingly, Kılıç and Şen (2014) asserted that three dimensions make up critical thinking disposition, innovation, participation, and cognitive maturity. Innovation refers to continuous pursuit of new ideas, openness to change, and a willingness to investigate developments in the educational environment. Participation is defined as the willingness of individuals to solve problems as they are encountered, while cognitive maturity indicates making evaluations based upon scientific data when exercising judgement. Thus, teachers with strong critical thinking skills, as they have necessary abilities to analyze, reflect, and make sound decisions, are more likely to promote autonomy in educational settings. Empirical research conducted with students supports this view by showing that critical thinking disposition has a moderate to strong correlation with teachers' autonomy support (Bolaños-Medina & Núñez, 2022; Koçoğlu & Kanadlı, 2019; Sevari & Farzadi, 2022). Therefore, it is predicted that teachers' critical thinking tendency will affect their autonomy support. Thus, we hypothesize that:

Hypothesis 1: Critical thinking disposition is positively and directly related to teachers' autonomy support.

Teacher Autonomy as a Mediator Between Critical Thinking Disposition and Teachers' Autonomy Support

Teacher autonomy refers to the teachers' ability and power to make their own decisions concerning teaching environment and school-wide matters (Çolak, 2025b; Strong & Yoshida, 2014). The term also represents the freedom of teachers to develop a unique pedagogy that requires harmony between training, personality, experience, and the requirements of a particular educational setting (Hoyle & John, 1995). Such freedom is important for both a school's development and fostering teachers' professionalism and the professional growth, as well as in solving school-related problems (Çolak 2025a; Çolak & Altınkurt, 2017). Furthermore, teacher autonomy acts as a cornerstone in helping students gain autonomy (Jiang & Ma, 2012). Thus, the literature points to the dependence of teachers' autonomy support on their own autonomous decision making. For example, in a study by Yazıcı (2016), teacher autonomy was proved to be a significant predictor of teacher autonomy support, whilst Basri

(2023) explored the factors affecting learner autonomy in a study conducted with teacher educators and revealed teacher autonomy as a facilitator for learner autonomy support. On the other hand, when teachers have limited autonomy over their professional activities and learning environment, they may provide less autonomy support to their learners, thereby exhibiting more controlling behaviors during the teaching process. For example, administrative pressures and accountability standards on curriculum or assessment may compel teachers to adopt a controlling pedagogical approach, which may result in teachers' strict conformity to regulations (Pelletier et al., 2002). Thus, teacher autonomy may serve as a strong predictor of teachers' autonomy support, such that teachers with greater discretion are more likely to promote students' autonomy in the learning process, whereas those with restricted autonomy may exhibit more controlling behaviors toward students.

Teacher autonomy has also close links with teachers' critical thinking disposition. Critical thinking supports the reasoning abilities of those who are independent thinkers (Cuypers, 2004; Kamii, 1991). Therefore, critical thinking may help enable teachers to exercise discretion within the processes of their profession. The literature underscores the existence of positive connections between critical thinking disposition and teacher autonomy. For example, Çolak et al. (2022) found moderate and positive links between critical thinking and teacher autonomy, with their study revealing that such a thinking style plays a crucial role in teachers' ability to exercise discretionary judgment. In addition, Lu and Wang (2021) concluded that teachers with more teaching, curriculum, and assessment autonomy are better able to utilize instruction that emphasizes critical thinking. These research examples highlight the interdependence established between teacher autonomy and critical thinking in the context of educational environments. Therefore, it can be asserted that higher levels of critical thinking disposition enable teachers to make independent decisions and display autonomous behaviors.

As has been seen, teacher autonomy may serve as both a predictor of teachers' autonomy support and an outcome of critical thinking, thereby suggesting its potential mediating role between these variables. It may be predicted that teachers having critical thinking skills such as being able to interpret and make inferences are more likely to exercise autonomy and make independent decisions. With higher autonomy, teachers may be better able to support their students' independence by creating a more self-directed learning environment. In other words, teachers with high critical thinking tendencies tend to exercise greater autonomy in their educational practices, and this in turn, may increase their level of support for learner autonomy. Thus, we hypothesize that:

Hypothesis 2: Teacher autonomy mediates the relationship between critical thinking and teachers' autonomy support.

Moderating Role of Teachers' Self-Efficacy Beliefs

Self-efficacy, initially introduced by Bandura, is conceptualized as an individual's belief in their ability to effectively complete tasks using the resources available, even when faced with challenging situations (Bandura, 1977). Self-efficacy beliefs impact the quality of human functioning across motivational, cognitive, emotional, and decision-making domains (Burić & Kim, 2020). Furthermore, self-efficacy beliefs have a domain-specific nature and manifest diversely according to the field of activity and specific context (Bandura, 2012). On the other hand, self-efficacy belief in teaching indicates a teacher's confidence in their own ability to plan, implement, and manage instruction effectively within their subject area (Burić & Kim, 2020).

Teachers' self-efficacy beliefs can impact upon their teaching processes as well as the learning processes, attitudes, achievements, and performance of their students (Tschannen-Moran & Hoy, 2007). This link is important since the teaching quality levels of teachers relate to their students' learning and achievement level (Hattie, 2009). Teachers with high self-efficacy beliefs tend to persist in the face of challenges, demonstrate greater perseverance, and are thereby generally more likely to achieve successful outcomes (Linnenbrink & Pintrich, 2003; Pajares, 1996). In addition, self-efficacy is closely associated with individuals' thinking skills. For example, higher levels of self-efficacy have been found to be positively related to the use of deep processing strategies, which involve analysis, reflection, and

integration of new knowledge (Fenollar et al., 2007). Other studies have shown clear links between teachers' self-efficacy and both creative and critical thinking (Orakci & Durnali, 2023; Zangenehvandi et al., 2014). Thus, teachers' confidence in their abilities may enable them to more effectively reflect their critical thinking skills into educational decisions and practices. In other words, teachers' higher levels of self-efficacy may strengthen the association between their critical thinking disposition and the degree of autonomy they display. We therefore put forward the following hypothesis:

Hypothesis 3: Teacher self-efficacy moderates the link between critical thinking disposition and teacher autonomy, such that the strength of the relationship is higher for individuals with high self-efficacy compared to those with low self-efficacy.

The links between self-efficacy and autonomy can be explored within the framework of social cognitive theory. According to the tenets of this theory, self-efficacy is recognized as a pivotal factor for human agency (Bandura, 1986) which acts as a positively motivating psychological mechanism (Stajkovic & Luthans, 1998). In addition, as Bandura (2012) suggested, teachers' self-efficacy beliefs can affect the choices they make and thereby shape their professional life. A systematic review conducted by Çolak (2024) points out that self-efficacy, together with professionalism, professional development, and leadership, constitutes a major component of teacher autonomy. Accordingly, choices that teachers make freely, and the initiatives and decisions they take, will increase their level of autonomy during the teaching process. Therefore, it can be concluded that teachers with high self-efficacy beliefs are more inclined to act autonomously in their educational practices. Empirical research in the literature have pointed out clear links between these variables. For example, Çolak (2025b) found that self-efficacy directly and positively influenced all dimensions of teacher autonomy, including teaching, curriculum, professional development, and communicative autonomy. Other studies have also revealed positive correlations between teachers' self-efficacy and their autonomy (Skaalvik & Skaalvik, 2014; Yukselir & Ozer, 2022). Thus, greater self-efficacy may facilitate teachers in effectively demonstrating autonomy in their professional practices. This influence may promote the impact of critical thinking on teachers' autonomy support through the mediating role of teacher autonomy. On the other hand, lower self-efficacy may constrain teachers' ability to take initiatives and make their own decisions, thereby reducing the mediated effect of autonomy. That is, self-efficacy may conditionally influence the strength of the indirect effect of critical thinking disposition on autonomy support via teacher autonomy, implying that an indirect effect via teacher autonomy becomes stronger with high self-efficacy. Thus, the final hypothesis of the study is as follows:

Hypothesis 4: Teacher self-efficacy moderates the strength of the mediated link between critical thinking disposition and teachers' autonomy support, such that the mediated relationship is stronger when teachers' self-efficacy is higher.

Method

Study Design

This research, which examines the role of teacher autonomy and self-efficacy in the relationship between teachers' critical thinking disposition and autonomy support, was designed in the survey model.

Participants

The population of the research consisted of 13,877 teachers working in elementary, lower secondary, and upper secondary schools in the Muğla province of Türkiye during the 2021-2022 academic year. The study's sample was chosen using the disproportionate cluster sampling technique. The number of samples capable of representing the population was determined as being at least 374 at a 95% confidence level. Measurement tools were applied based on the voluntary participation of teachers. Within the context of the study, data were gathered from a total of 440 teachers. Of the data initially obtained, 23 sets of data (16 one-way, 7 multi-directional) were removed since they were assessed to represent extreme values. Therefore, the study's analysis was conducted with 417 valid sets of data. Of the teachers participating in the research, 62.4% identified as female (n = 260) whilst 37.6%

(n = 157) identified as male. In terms of the school type where the participants worked, 36.9% were from elementary schools (n = 154), 30% from lower secondary schools (n = 125), and 33.1% worked in upper secondary schools (n = 138). From examining the teachers' seniority (length of teaching service), 15.8% (n = 66) had 9 years of seniority or less, 36.9% (n = 154) had 10-19 years of seniority, and 47.2% (n = 197) had 20 years or more of seniority.

Data Collection Measures

Prior to collecting the data, ethics approval was granted by the Social and Human Sciences Research Ethics Committee of Muğla Sıtkı Koçman University (Decision number 39, dated 10 March 2022), as well as institutional permission granted by the Muğla Provincial Directorate of National Education. The UF/EMI Critical Thinking Disposition Scale, Teacher Self-Efficacy Beliefs Scale, Teacher Autonomy Scale, and Learner Autonomy Support Scale were utilized as data collection tools in the study.

Critical Thinking Disposition Scale. The UF/EMI Critical Thinking Disposition Scale, originally developed by Irani et al. (2007), was adapted to the Turkish context by Kılıç and Şen (2014). The scale consists of 25 Likert-type items, with responses in a range from 1 (*Strongly disagree*) to 5 (*Strongly agree*) and is structured into three factors: Participation, Cognitive Maturity, and Innovativeness (Kılıç & Şen, 2014). In the current study, Cronbach's alpha internal consistency coefficients were determined to be .85 for the Participation dimension, .74 for the Cognitive Maturity dimension, .75 for the Innovativeness dimension, and .92 for the overall total score of the scale. We checked the construct validity of the scale using CFA (Confirmatory Factor Analysis) which indicated an acceptable fit: χ^2 / df = 3.99, RMSEA = .09, RMR = .03, SRMR = .06, CFI = .87, GFI = .85, and IFI = .87 (Schermelleh-Engel et al., 2003). We also checked validity using average variance extracted (AVE) results (AVE = 0.59), which indicated that the observed variables adequately represented the construct.

Teacher Self-Efficacy Beliefs Scale. The current research also employed the Teacher Self-efficacy Beliefs Scale developed by Çolak et al. (2017). The scale consists of 27 items and four factors, with scale item responses in a range from 1 (*Disagree*) to 5 (*Agree*). These four factors are Intellectual Self-Efficacy, Professional Self-Efficacy, Academic Self-Efficacy, and Social Self-Efficacy. Within the framework of the current research, we concluded that the Cronbach's alpha internal consistency coefficients were .88 for the Professional Self-Efficacy dimension, .82 for the Academic Self-Efficacy dimension, .88 for the Intellectual Self-Efficacy dimension, .87 for the Social Self-Efficacy dimension, and .93 for the entire scale. We checked the construct validity of the scale using CFA which indicated a good fit: χ^2 / df = 2.87, RMSEA = .07, RMR = .03, SRMR = .07, CFI = .90, GFI = .85, and IFI = .90 (Schermelleh-Engel et al., 2003). We also checked validity using average variance extracted (AVE) results (AVE = 0.66), which indicated that the observed variables adequately represented the construct.

Teacher Autonomy Scale. In the study, we measured teacher autonomy using the Teacher Autonomy Scale developed by Çolak and Altınkurt (2017). There are 17 items in the scale, with item responses in a range from 1 (*Strongly disagree*) to 5 (*Strongly agree*). The scale consists of four factors: Teaching Autonomy, Curriculum Autonomy, Communicative Autonomy, and Professional Development Autonomy (Çolak & Altınkurt, 2017). In the current study, Cronbach's alpha internal consistency coefficients were determined as being .71 for the Teaching Autonomy, .79 for the Curriculum Autonomy, .69 for the Communicative Autonomy, and .73 for the Professional Development Autonomy, and .83 for the overall total score of the scale. We checked the construct validity of the scale using CFA which indicated an acceptable fit: χ^2 / df = 3.27, RMSEA = .07, RMR = .06, SRMR = .08, CFI = .90, IFI = .90, and GFI = .91 (Schermelleh-Engel et al., 2003). We also checked validity using average variance extracted (AVE) results (AVE = 0.50), which indicated that the observed variables adequately represented the construct.

Learner Autonomy Support Scale. This scale was developed by Oğuz (2013) to measure teachers' autonomy support. The scale comprises 16 Likert-type items, with responses in a range from 1 (*Never*) to 5 (*Always*). The scale consists of two subscales: necessity and display of behaviors (Oğuz,

2013). For the current research, the display subscale was used. This subscale includes three dimensions: Support for Learning Process, Support for Feelings and Thoughts, and Support for Assessment. Cronbach's alpha internal consistency coefficients were determined to vary between .84 and .87 in the sub-dimensions and are .93 for the overall display scale. We checked the construct validity of the scale using CFA which indicated a good fit: χ^2 / df = 2.72, RMSEA = .06, RMR = .02, SRMR = .04, CFI = .96, GFI = .93, and IFI = .96 (Schermelleh-Engel et al., 2003). We also checked validity using average variance extracted (AVE) results (AVE = 0.72), which indicated that the observed variables adequately represented the construct.

Data Analysis

First, we performed CFA in order to determine the construct validity of the four scales used in the study's data collection. IBM's AMOS 22 software was used for this analysis. Then we carried out descriptive statistical analysis and tested correlations between the study's variables. The mediation and moderation hypotheses of the study were examined using a macro (named PROCESS) that was developed by Hayes (2013) for IBM's SPSS. Hayes' Model 4 was then utilized to test the mediation (Hypothesis 2) along with the direct effect of critical thinking disposition on teachers' autonomy support (Hypothesis 1). Hayes' Model 7 was employed for testing moderation in the study. The moderation of an effect either to or from a mediator can be tested by this model (Hayes, 2013). Model 7, which was described as a first-stage moderation model by Edwards and Lambert (2007), enabled us to test both Hypothesis 3 as the moderator hypothesis and Hypothesis 4 as the moderated mediation hypothesis. While conducting our analyses, we used a 5,000 bootstrap count and 95% confidence interval (CI), as well as opting for mean-centering when estimating the effect.

Prior to conducting the data analysis, the collected data were examined in terms of the requirements for performing multivariate analysis. In this respect, we analyzed the data regarding single and multiple outliers, as well as for the problems of normality and multicollinearity. Having removed 23 outliers from the dataset, the study's analysis was subsequently conducted using data from a total of completed 417 scales. We examined kurtosis and skewness coefficients in order to assess the normality assumption of the variables in the study. The examination indicated that the skewness and kurtosis coefficients fell within the range of -1 to +1 (skewness -.48 to -.24; kurtosis -.83 to .26), which indicated normality of the distribution. Regarding multicollinearity, we calculated tolerance index (Tl > .10) and variance inflation factor (VIF) values (VIF < 10), which both demonstrated that no multicollinearity problem existed (Kline, 2016) among the variables. In the data set, the smallest tolerance value was .41, and the largest VIF value was 2.46.

Results

Descriptive Analysis

Descriptive statistics, scale reliabilities, and scale correlations are presented in Table 1. The average scores indicated that the participant teachers had high levels of critical thinking disposition (M = 4.32, SD = .45), self-efficacy beliefs (M = 4.27, SD = .43), and autonomy support (M = 4.40, SD = .47), whilst they demonstrated autonomy at almost a high level (M = 4.15, SD = .48). Cronbach's alpha coefficients indicated strong internal consistency for the scales. Correlations among the variables prove that critical thinking relates positively and significantly to teacher autonomy (r = 0.51, p < .01), self-efficacy (r = 0.75, p < .01), and teachers' autonomy support (r = 0.53, p < .01). The results further indicate that teacher autonomy correlates positively and significantly to self-efficacy (r = 0.53, p < .01) and teachers' autonomy support (r = 0.68, p < .01).

Table 1. Descriptive analysis results

Variable	*M*	*SD*	*α*	1	2	3	4
1. Critical thinking disposition	4.32	.45	.95	-			
2. Teacher autonomy	4.15	.48	.85	.51**	-		
3. Self-efficacy	4.27	.43	.93	.75**	.53**	-	
4. Teachers' autonomy support	4.40	.47	.93	.53**	.68**	.56**	-
Skewness	-	-	-	-.24	-.43	-.38	-.48
Kurtosis	-	-	-	-.83	.26	-.58	-.63
Tl	-	-	-	.42	.69	.41	-
VIF	-	-	-	2.39	1.44	2.46	-

Note. *: p < .05, ** p < .01, M: Mean, SD: Standard deviation, α: Cronbach's alpha coefficient, VIF: Variance inflation factor, Tl: Tolerance index.

Validity Analysis

We utilized IBM's AMOS 22 software to perform CFA as part of our analysis to evaluate the validity of the four scales employed in the study. The hypothesized four-factor model demonstrated a strong fit with the data, χ^2 = 186.61, df = 71, RMSEA = .06, RMR = .01, SRMR = .04, CFI = .97, GFI = .94, and IFI = .97 (Hu & Bentler, 1999). We also observed that the hypothesized four-factor model exhibited a significantly better fit with the data compared to the three-factor model where self-efficacy and teacher autonomy were combined into a single factor, χ^2 = 353.68, df = 74, RMSEA = .10, RMR = .02, SRMR = .06, CFI = .92, GFI = .88, and IFI = .92. The examination of factor loadings revealed that all indicators in the model significantly loaded onto their respective factor with standardized coefficients falling between .53 and .97 (p < .001).

Mediation Analysis

Initially, we found that critical thinking disposition had a positive and significant effect on teacher autonomy support (b = .25, SE = .04, 95% CI [.17, .34]), which confirmed Hypothesis 1. Our second hypothesis predicted that teacher autonomy mediated the relationship between critical thinking disposition and autonomy support. We observed that teacher autonomy had a significant, positive, and direct effect on autonomy support (b = .53, SE = .04, 95% CI [.46, .62]). In terms of mediation, as depicted in Table 2, indirect effects of critical thinking on teacher autonomy support via teacher autonomy were significant and positive (b = .28, SE = .03, 95% CI [.23, .34]). Thus, Hypothesis 2 was supported.

Moderation Analysis

Our third hypothesis predicted that teacher self-efficacy moderated the link between critical thinking disposition and teacher autonomy. Analysis revealed that the interaction between critical thinking and self-efficacy was significant when predicting teacher autonomy (b = .22, SE = .10, 95% CI [.03, .42]), which confirmed Hypothesis 3. To further understand the nature of the moderation, simple slope analysis was performed (Aiken & West, 1991) which revealed that the positive relationship between critical thinking and teacher autonomy was stronger (simple slope = .37, p < .01) under high self-efficacy (+1 SD) but weaker (simple slope = .18, p < .05) under low self-efficacy (-1 SD) (see Figure 2).

Table 2. Findings on mediation and moderation effects

	Outcome Variable: Teachers' Autonomy Support				
				Bootstrapped 95% CI	
Predicting variables	**Coefficient**	***SE***	***p***	**LL**	**UL**
	Mediation model (Model 4)				
Teacher autonomy	.53	.04	<.01	.46	.62
Critical thinking disposition	.25	.04	<.01	.17	.34
Indirect effect via teacher autonomy	.28	.03	<.01	.23	.34
	Outcome Variable: Teacher Autonomy				
	Moderated mediation model (Model 7)				
Critical thinking disposition	.27	.07	<.01	.14	.40
Self-efficacy	.40	.07	<.01	.26	.53
Critical thinking disposition x Self-efficacy	.22	.10	<.01	.03	.42

Note. N = 417, CI = Confidence interval.

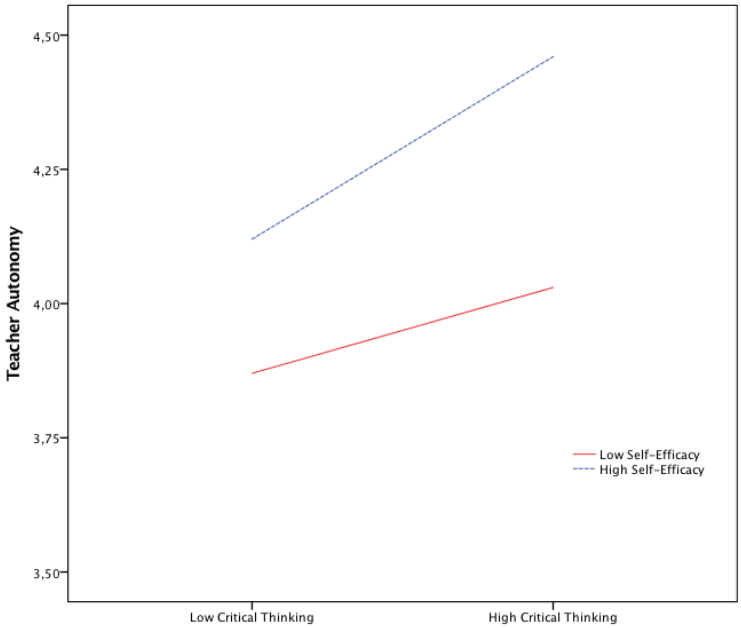


Figure 2. Effect of self-efficacy on relationship between critical thinking and teacher autonomy

Moderated Mediation Analysis

Our fourth hypothesis further predicted that teacher self-efficacy moderated the mediated relationship between critical thinking disposition and teachers' autonomy support. The indirect effects of critical thinking on teachers' teacher autonomy support (b = .20, 95% CI [.11, .28]) via teacher autonomy were found to be stronger for those exhibiting a high level of self-efficacy (see Table 3). Nevertheless, the indirect effects of critical thinking on teachers' autonomy support (b = .09, 95% CI [.01, .18]) were determined to be weaker for teachers who exhibited a low level of self-efficacy. Furthermore, when predicting teachers' autonomy support, the index of moderated mediation also proved to be significant (index = .12, 95% CI [.01, .22]). Thus, Hypothesis 4 was confirmed.

Table 3. Conditional indirect effect of critical thinking on autonomy support via teacher autonomy at levels of self-efficacy (PROCESS, Model 7)

			Bootstrapped 95% CI	
Conditional indirect effects (via teacher autonomy)	**Coefficient**	***SE***	**LL**	**UL**
	Critical thinking to teachers' autonomy support			
Low self-efficacy (-1 *SD*)	.09	.05	.01	.18
High self-efficacy (+1 *SD*)	.20	.04	.11	.28
Index of moderated mediation	.12	.05	.01	.22

Discussion

While earlier research conducted with students has explored the connections between critical thinking disposition and autonomy support (Bolaños-Medina & Núñez, 2022; Koçoğlu & Kanadlı, 2019; Sevari & Farzadi, 2022), there exists a generally limited understanding of teachers' critical thinking disposition as a predictor of their autonomy support. In addition, no prior study has examined the possible mechanism that explains the association between these variables. In the current study, we examined the mediating role of teacher autonomy and the moderating role of teachers' self-efficacy beliefs in the relationship between teachers' critical thinking disposition and autonomy support along with the moderating function of teachers' self-efficacy beliefs in the relationship between critical thinking disposition and teacher autonomy. We determined that teacher autonomy mediated the links between critical thinking disposition and teachers' autonomy support, with self-efficacy moderating the mediated effect of teacher autonomy in this relationship. Furthermore, we concluded that self-efficacy was a significant moderator in the link between critical thinking disposition and teacher autonomy.

Our results illustrate the significant relationship between critical thinking disposition and teachers' autonomy support. Teachers with higher levels of critical thinking disposition are more likely to display autonomy support for their students. Although to the best of our knowledge, the literature has not directly examined the effect of teachers' critical thinking disposition on their autonomy support, some studies have found a moderate to strong correlation between students' critical thinking disposition and autonomy support (Bolaños-Medina & Núñez, 2022; Koçoğlu & Kanadlı, 2019; Sevari & Farzadi, 2022). Concerning teachers' thinking skills, a study conducted by Orakci and Durnali (2023) showed that teachers' creative thinking has a positive and strong relationship with their autonomy support. Higher-order thinking abilities, including critical thinking involves a decision-making process that includes analysis, interpretation, evaluation, and inference (Çolak et al., 2019). These skills may enable teachers to interpret, make independent decisions, search for innovations, and be more readily accepting of change (Kılıç & Şen, 2014). In addition, critical thinking abilities have been shown to be essential for teachers' effectiveness (Birjandi & Bagherkazemi, 2010) and professional development (Liao et al., 2022), both of which may encourage teachers to assist students in making their own decisions. Thus, as suggested by the current study, teachers with higher critical thinking disposition tend to provide more support for students' autonomy by identifying their needs and permitting their interests and preferences to direct classroom activities.

Another contribution of the current research to the literature lies in confirming the mediating effect of teacher autonomy in the relationship between critical thinking disposition and teachers' autonomy support. This indicates that critical thinking disposition promotes teachers' autonomy support through the enhancement of their autonomy. The mediating effect arises from teacher autonomy being linked to both critical thinking and autonomy support. This finding corresponds with the results of earlier studies which established a positive relationship between teachers' critical thinking disposition and their autonomy (Çolak et al., 2022; Lu & Wang, 2021). Critical thinking generally leads to a higher level of reasoning (Kamii, 1991), thereby enabling teachers to make independent yet

informed decisions. Some studies have shown positive outcomes of critical thinking for educators in areas such as teaching success (Nosratinia & Zaker, 2017) and reflective thinking (Erdoğan, 2020). In addition, regarding teacher candidates, Tunçeli et al. (2022) concluded that critical thinking promoted their autonomy levels, explaining nearly one-third of the variance in leaners' autonomous behaviors. The literature further suggests a positive and significant link between teacher autonomy and teachers' autonomy support (Yazıcı, 2016). Drawing on the perspectives of academics and university students, Basri (2023) found teacher autonomy as a facilitator for learner autonomy support, indicating that when teachers exercise increased autonomy in their decision making, they also tend to be supportive of their students' autonomy. Pelletier et al. (2002) showed that the more teachers are self-determined in their work, the more they demonstrate supportive behaviors toward students. As evidenced in the literature, teacher autonomy is closely associated with both critical thinking and autonomy support. The current study extends upon this by having explored teacher autonomy as an essential mechanism through which critical thinking disposition positively affects teachers' autonomy support. Thus, teacher autonomy is crucial for conveying the positive impact of critical thinking on the autonomy supportive actions displayed by teachers.

Our results also revealed teacher self-efficacy as a significant moderating role in the relationship between critical thinking disposition and teacher autonomy. This suggests that self-efficacy strengthens the effect of critical thinking disposition on teacher autonomy; that is, teachers with high self-efficacy beliefs are more inclined to exhibit critical thinking disposition, consequently demonstrating greater autonomous behaviors. In addition to indicating confidence in one's abilities, self-efficacy beliefs also influence individuals' cognitive, emotional, and decision-making processes (Burić & Kim, 2020), which emphasizes the key role of self-efficacy in developing thinking skills. For example, Orakci and Durnali (2023) concluded that teachers' self-efficacy is significantly and positively correlated with their creative thinking, while Zangenehvandi et al. (2014) found significant and positive correlation between teachers' self-efficacy and critical thinking. Regarding university students, Phan (2009) found that academic self-efficacy influenced critical thinking positively by making students more inclined to participate in learning activities that aim to challenge their presumptions, beliefs, and awareness. In addition, higher self-efficacy beliefs have been shown to positively linked to deep processing strategies, including reflection, analysis, and incorporation of new knowledge (Fenollar et al., 2007). When strengthened by self-efficacy, higher critical thinking levels lead teachers to exhibit greater autonomy. Therefore, it may be concluded that teachers will exercise higher autonomy when they have strong belief in their own capabilities and are convinced that their skills will enable them to attain favorable outcomes.

Finally, our study highlights that self-efficacy moderates the strength of the mediated relationship between critical thinking disposition and teachers' autonomy support. This implies that when teachers experience higher levels of self-efficacy, it fosters an increased sense of higher professional autonomy, and which may ultimately lead to increased autonomy support for their learners. In contrast, for teachers with low self-efficacy, the mediation effect of teacher autonomy in the link between critical thinking disposition and autonomy support becomes weaker. Our finding corresponds with the results of earlier studies examining the connection between teachers' self-efficacy and autonomy (Çolak, 2025b; Skaalvik & Skaalvik, 2014). For example, Sokmen and Kilic (2019) established self-efficacy as a positive predictor of teacher autonomy, and similarly, Lu et al. (2015) revealed significant and positive links between self-efficacy and teacher autonomy. Kasapoğlu Tankutay and Çolak (2025) concluded that academic optimism, with self-efficacy recognized as one of its key components, enable teachers to make their own decisions by promoting their autonomy. The current study's results also align with the principals of social cognitive theory, which asserts that self-efficacy serves as the underlying psychological mechanism that provides positive motivation for individuals (Stajkovic & Luthans, 1998). According to this theory, in order to activate human agency and persuade people to take action, they should be sufficiently convinced that they are able to reach the

desired results (Bandura, 1986). The theory suggests that self-efficacy can be considered as a motivating mechanism that triggers individuals to take action and influences their choices. As shown in this study, higher levels of self-efficacy lead teachers to provide greater autonomy support to their students by enhancing their own autonomy.

Overall, our study suggests that self-efficacy together with critical thinking play a crucial role in guiding teachers to make their own decisions, thereby fostering increased support for their students' autonomous behaviors. Teachers' autonomy support is of crucial importance in educational settings, as it offers considerable benefits for both student development and instructional outcomes. In their experimental study, Reeve et al. (2004) found that teachers' autonomy support improved students' engagement in the teaching process, which serves as a key indicator of their underlying motivation. Some studies have shown the interconnection between teachers' autonomy support and student autonomy, indicating that greater support form teachers enabled students to make independent decisions and act accordingly (Basri, 2023; Núñez et al., 2015). Other studies have reported that such support contributes to students' well-being and satisfaction (Black & Deci, 2000; Sheldon & Krieger, 2007). Most importantly, a longitudinal study by Fu et al. (2023) concluded that teachers' autonomy support positively and directly influenced students' academic performance both immediately and over time. Therefore, in order to benefit from teachers' autonomy supportive behaviors in educational environments, it is essential to uncover their underlying mechanisms. This study provided evidence for how the improvement of critical thinking, along with self-efficacy and teacher autonomy, contributes to the development of autonomy supportive teachers.

Practical Implications

Our study results have several practical implications. First, our study shows that critical thinking disposition positively affect both teacher autonomy and teachers' autonomy support. Critical thinking refers to a cognitive process during which individuals systematically interpret, analyze, and evaluate information in order to make their own reasoned judgements and decisions (Facione, 1990). Therefore, it can be inferred that the improvement of critical thinking skills requires a relatively long period of time to cultivate. In order to develop the critical thinking skills of teachers, preservice teacher training programs should be structured in such a way that they offer significant possibilities and opportunities for teacher candidates to apply their knowledge during the processes of decision making and problem solving. In addition, students' interest in acquiring new knowledge will help motivate them to relate it to their prior knowledge, fostering critical thinking in the process (Bolaños-Medina & Núñez, 2022). Thus, teacher candidates should also be motivated by being exposed to engaging activities and effective learning methods in order to acquire new knowledge during the education process.

Second, the mediating role of teacher autonomy proposes that teachers' autonomy support is partly affected by their own levels of autonomy. Türkiye has a highly centralized education system, implying that the autonomy levels of teachers in domains like shaping course content, curricula development, and influencing school-related policies are significantly lower than the OECD average (Schleicher, 2020). Within such a system, teachers have limited opportunity to exercise discretion in their professional activities. In order to enhance teacher autonomy in countries such as Türkiye, national-level regulatory changes are needed that grant teachers the right to organize their own teaching and to tailor the curriculum they teach as required. In addition, as a positive school environment fosters the development of teacher autonomy (Çolak & Altınkurt, 2017), we suggest that school principals should endeavor to establish and cultivate a motivating and supportive school environment, since this would enhance teacher autonomy, which in turn would promote autonomy support provided to the students.

Third, our study concluded that the indirect effect of critical thinking on autonomy support is contingent upon self-efficacy belief levels. Thus, there is a need for teachers' self-efficacy beliefs to be developed and strengthened. The principles of social cognitive theory may serve as a guide for how teachers' self-efficacy beliefs can be increased. Bandura (1997) asserted that mastery experience constitutes an important component of an individual's self-efficacy belief. In this regard, school principals should create appropriate opportunities for teachers to develop their professional abilities and thereby enhance the instructional quality offered at their school. We recommend that school principals support teachers' professional learning by helping to motivate them to attend seminars as well as setting school-wide instructional objectives. This approach will support teachers to accomplish challenging tasks and develop mastery skills, which in turn will develop their self-efficacy beliefs.

Limitations and Future Research

Despite having implications at practical level, it is crucial to recognize that the current study presents certain limitations. First, this study relied upon a teacher sample from one province in Türkiye; hence, future research could consider expanding the sample source to enhance a wider representation of the population or through drawing samples from different countries in order to compare the current study's results with diverse cultural environments. Second, the study was conducted across three different school levels, which limits the generalizability of the results. Third, the data of the study originated from a single source, teachers' self-reports, which may have excluded the perspectives of students, administrators, and trained raters. Thus, forthcoming studies could consider including different sources of data collection in order to bolster the credibility and validity of the results gained from the current study. Fourth, this study's focus was on critical thinking disposition as a single construct rather than analyzing its dimensions separately. We recommend that future studies investigate the multifaceted nature of critical thinking disposition by considering its various dimensions, such as participation, cognitive maturity, and innovativeness. Finally, this study explored the mediating effect of teacher autonomy between the study variables. Future research could also investigate other possible mediators such as academic optimism, teacher innovativeness, subjective well-being, and teacher professional learning

References

Aiken, L. S., & West, S. G. (1991). *Multiple regression: Testing and interpreting interactions.* Sage.

Akçil, M., & Oğuz, A. (2015). Investigation of the relationship between the science teachers' self-efficacy beliefs and their learner autonomy support behaviours. *Electronic Turkish Studies*, *10*(11), 1-16. http://doi.org/10.7827/TurkishStudies.8616

Alkın-Şahin, S., Tunca, N., & Oğuz, A. (2015). Classroom teachers' supportive behaviors for learner autonomy and critical thinking. *Route Educational and Social Science Journal*, *2*(1), 161-178.

Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review, 84*(2), 191-215. https://doi.org/10.1037/0033-295X.84.2.191

Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory.* Prentice-Hall.

Bandura, A. (1997). *Self-efficacy: The exercise of control.* Freeman.

Bandura, A. (2012). Cultivate self-efficacy for personal and organizational effectiveness. In E. A. Locke (Ed.), *Handbook of principles of organizational behavior: Indispensable knowledge for evidence-based management* (pp. 179-200). John Wiley & Sons. https://doi.org/10.1002/9781119206422.ch10

Basri, F. (2023). Factors influencing learner autonomy and autonomy support in a faculty of education. *Teaching in Higher Education, 28*(2), 270-285. https://doi.org/10.1080/13562517.2020.1798921

Birjandi, P., & Bagherkazemi, M. (2010). The relationship between Iranian EFL teachers' critical thinking ability and their professional success. *English Language Teaching, 3*(2), 135-145. https://doi.org/10.5539/elt.v3n2p135

Black, A. E., & Deci, E. L. (2000). The effects of student self-regulation and instructor autonomy support on learning in a college-level natural science course: A self-determination theory perspective. *Science Education, 84*(6), 740-756. https://doi.org/10.1002/1098-237X(200011)84:6%3C740::AID-SCE4%3E3.0.CO;2-3

Bolaños-Medina, A., & Núñez, J. L. (2022). Autonomy support, critical thinking, and motivation as key predictors of translator trainees' strategic competence. *Across Languages and Cultures, 23*(2), 131-147. https://doi.org/10.1556/084.2022.00266

Burić, I., & Kim, L. E. (2020). Teacher self-efficacy, instructional quality, and student motivational beliefs: An analysis using multilevel structural equation modeling. *Learning and Instruction,* (66), 101302. https://doi.org/10.1016/j.learninstruc.2019.101302

Cuypers, S. E. (2004). Critical thinking, autonomy and practical reason. *Journal of Philosophy of Education, 38*(1), 75-90. https://doi.org/10.1111/j.0309-8249.2004.00364.x

Çolak, İ. (2024). Science mapping of the knowledge base on teacher autonomy: A bibliometric study. *International Journal of Educational Management, 38*(5), 1305–1324. https://doi.org/10.1108/IJEM-05-2023-0241

Çolak, İ. (2025a). The evolving knowledge base on teacher professionalism: a bibliometric review from 1960 to 2024. *European Journal of Teacher Education*, 1-24. https://doi.org/10.1080/02619768.2025.2554691

Çolak, İ. (2025b). The relationship between organizational trust and teacher autonomy: The mediating role of self-efficacy beliefs. *Psychology in the Schools*, *62*(5), 1539-1549. https://doi.org/10.1002/pits.23414

Çolak, İ., & Altınkurt, Y. (2017). The relationship between school climate and teacher autonomy behaviors. *Educational Administration: Theory and Practice, 23*(1), 33-71. https://doi.org/10.14527/kuey.2017.002

Çolak, İ., Yorulmaz, Y. İ., & Altınkurt, Y. (2017). The validity and reliability study of teacher self-efficacy beliefs scale. *MSKU Journal of Education, 4*(1), 20-32. https://doi.org/10.21666/muefd.319209

Çolak, İ., Yorulmaz, Y. İ., & Altınkurt, Y. (2022). The mediating role of teachers' critical thinking dispositions in the relationship between their organisational power distance perceptions and autonomy behaviors. *E-International Journal of Educational Research, 13*(2), 90-107. https://doi.org/10.19160/e-ijer.1076433

Çolak, İ., Anasız, B. T., Yorulmaz, Y. İ., & Duman, A. (2019). The role of gender, grade level, and parents' educational attainment variables on teacher candidates' critical thinking dispositions: A meta-analysis. *E-International Journal of Educational Research, 10*(1), 67-86. https://doi.org/10.19160/ijer.541861

Deci, E. L., & Ryan, R. M. (1991). A motivational approach to self: Integration in personality. In R. A. Dienstbier (Ed.), *Perspectives on motivation* (pp. 237-288). University of Nebraska Press.

Deci, E. L., & Ryan, R. M. (2000). The" what" and" why" of goal pursuits: Human needs and the self-determination of behavior. *Psychological Inquiry*, *11*(4), 227-268. https://doi.org/10.1207/S15327965PLI1104_01

Edwards, J. R., & Lambert, L. S. (2007). Methods for integrating moderation and mediation: A general analytical framework using moderated path analysis. *Psychological Methods, 12*(1), 1-22. https://doi.org/10.1037/1082-989X.12.1.1

Erdoğan, F. (2020). The relationship between prospective middle school mathematics teachers' critical thinking skills and reflective thinking skills. *Participatory Educational Research, 7*(1), 220-241. https://doi.org/10.17275/per.20.13.7.1

European Union. (2019). *Key competences for lifelong learning.* https://op.europa.eu/en/publication-detail/-/publication/297a33c8-a1f3-11e9-9d01-01aa75ed71a1/language-en

Facione, P. A. (1990). *Critical thinking: A statement of expert consensus for purposes of educational assessment and instruction. research findings and recommendations.* American Philosophical Association.

Fenollar, P., Román, S., & Cuestas, P. J. (2007). University student's academic performance: An integrative conceptual framework and empirical analysis. *British Journal of Educational Psychology, 77*(4), 873-891. https://doi.org/10.1348/000709907X189118

Fu, D., Liu, Y., & Zhang, D. (2023). The relationship between teacher autonomy support and student mathematics achievement: A 3-year longitudinal study. *Educational Psychology*, *43*(2-3), 187-206. https://doi.org/10.1080/01443410.2023.2190064

Guay, F., Boggiano, A. K., & Vallerand, R. J. (2001). Autonomy support, intrinsic motivation, and perceived competence: Conceptual and empirical linkages. *Personality and Social Psychology Bulletin*, *27*(6), 643-650. https://doi.org/10.1177/0146167201276001

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

Hayes, A. F. (2013). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach*. Guilford Publications.

Hoyle, E., & John, P. (1995). *Professional knowledge and professional practice*. Cassell.

Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal, 6*(1), 1-55. https://doi.org/10.1080/10705519909540118

Huéscar Hernández, E., Moreno-Murcia, J. A., Cid, L., Monteiro, D., & Rodrigues, F. (2020). Passion or perseverance? The effect of perceived autonomy support and grit on academic performance in college students. *International Journal of Environmental Research and Public Health, 17*(6), 2143. https://doi.org/10.3390/ijerph17062143

Irani, T., Rudd, R., Gallo, M., Ricketts, J., Friedel, C., & Rhoades, E. (2007). *Critical thinking instrumentation manual.* University of Florida. https://www.tntech.edu/citl/pdf/critical-thinking/UF-EMI.pdf

Jang, H., Reeve, J., & Deci, E. L. (2010). Engaging students in learning activities: It is not autonomy support or structure but autonomy support and structure. *Journal of Educational Psychology*, *102*(3), 588-600. https://doi.org/10.1037/a0019682

Jiang, Y., & Ma, T. (2012). A comparative study of teacher autonomy between novice teachers and proficient teachers in the context of university English teaching reform in China. *Sino-US English Teaching, 9*(3), 963-974.

Kamii, C. (1991). Toward autonomy: The importance of critical thinking and choice making. *School Psychology Review, 20*(3), 382-388. https://doi.org/10.1080/02796015.1991.12085561

Kasapoğlu Tankutay, H., & Çolak, İ. (2025). How school principals' empowering leadership influences teacher autonomy: The mediating role of teachers' academic optimism. *British Educational Research Journal*, *51*(3), 1255-1270. https://doi.org/10.1002/berj.4125

Kılıç, H. E., & Şen, A. (2014). Turkish adaptation study of UF/EMI critical thinking disposition instrument. *Education and Science*, *39*(176), 1-12. https://doi.org/10.15390/EB.2014.3632

Kline, R. B. (2016). *Principles and practice of structural equation modeling* (4th ed.). The Guilford Press.

Koçoğlu, A., & Kanadlı, S. (2019). An investigation of secondary school students' perceived autonomy support, critical thinking tendencies and problem solving skills. *Trakya Journal of Education*, *9*(1), 61-77. https://doi.org/10.24315/trkefd.406627

Liao, W., Liu, M., Wang, Z., & Qin, K. (2022). Chinese expert teachers' critical thinking strategies for professional growth. *Professional Development in Education*, 1-15. https://doi.org/10.1080/19415257.2022.2097290

Linnenbrink, E. A., & Pintrich, P. R. (2003). The role of self-efficacy beliefs in student engagement and learning in the classroom. *Reading & Writing Quarterly, 19*(2), 119-137. https://doi.org/10.1080/10573560308223

Little, D. (1995). Learning as dialogue: The dependence of learner autonomy on teacher autonomy. *System, 23*(2), 175-181. https://doi.org/10.1016/0346-251X(95)00006-6

Lu, G., & Wang, J. (2021). Relationships between teacher autonomy, collaboration, and critical thinking focused instruction: A cross-national study. *International Journal of Educational Research,* (106), 101730. https://doi.org/10.1016/j.ijer.2020.101730

Lu, J., Jiang, X., Yu, H., & Li, D. (2015). Building collaborative structures for teachers' autonomy and self-efficacy: The mediating role of participative management and learning culture. *School Effectiveness and School Improvement, 26*(2), 240-257. https://doi.org/10.1080/09243453.2014.888086

Nosratinia, M., & Zaker, I. A. (2017). Scrutinizing the impact of teachers' critical thinking and teaching autonomy on their teaching success and learners' use of language learning strategies. *Journal of Language Teaching and Research, 8*(1), 122-132. https://doi.org/10.17507/jltr.0801.15

Núñez, J. L., Fernández, C., León, J., & Grijalvo, F. (2015). The relationship between teacher's autonomy support and students' autonomy and vitality. *Teachers and Teaching, 21*(2), 191-202. https://doi.org/10.1080/13540602.2014.928127

OECD. (2018). *The future of education and skills: Education 2030.* https://www.oecd.org/content/dam/oecd/en/publications/reports/2018/06/the-future-of-education-and-skills_5424dd26/54ac7020-en.pdf

Oğuz, A. (2013). Developing a scale for learner autonomy support. *Educational Sciences: Theory & Practice, 13*(4), 2187-2194. https://doi.org/10.12738/estp.2013.4.1870

Oğuz, A., Altınkurt, Y., Yılmaz, K., & Hatipoğlu, S. (2014). The relationship between educational beliefs and learner autonomy support behaviors of teachers. *Turkish Journal of Educational Studies, 1*(1), 37-78.

Orakci, Ş., & Durnali, M. (2023). The mediating effects of metacognition and creative thinking on the relationship between teachers' autonomy support and teachers' self-efficacy. *Psychology in the Schools*, *60*(1), 162-181. https://doi.org/10.1002/pits.22770

Pajares, F. (1996). Self-efficacy beliefs in academic settings. *Review of Educational Research, 66*(4), 543-578. https://doi.org/10.3102/00346543066004543

Pelletier, L. G., Séguin-Lévesque, C., & Legault, L. (2002). Pressure from above and pressure from below as determinants of teachers' motivation and teaching behaviors. *Journal of Educational Psychology*, *94*(1), 186-196. https://doi.org/10.1037//0022-0663.94.1.186

Phan, H. P. (2008). Predicting change in epistemological beliefs, reflective thinking and learn-ing styles: A longitudinal study. *British Journal of Educational Psychology, 78*(1), 75-93. https://doi.org/10.1348/000709907X204354

Phan, H. P. (2009). Relations between goals, self-efficacy, critical thinking and deep processing strategies: A path analysis. *Educational Psychology*, *29*(7), 777-799. https://doi.org/10.1080/01443410903289423

Reeve, J., & Cheon, S. H. (2016). Teachers become more autonomy supportive after they believe it is easy to do. *Psychology of Sport and Exercise*, (22), 178-189. https://doi.org/10.1016/j.psychsport.2015.08.001

Reeve, J., & Jang, H. (2006). What teachers say and do to support students' autonomy during a learning activity. *Journal of Educational Psychology, 98*(1), 209-218. https://doi.org/10.1037/0022-0663.98.1.209

Reeve, J., Jang, H., Carrell, D., Jeon, S., & Barch, J. (2004). Enhancing students' engagement by increasing teachers' autonomy support. *Motivation and Emotion,* (28), 147-169. https://doi.org/10.1023/B:MOEM.0000032312.95499.6f

Ryan, R., & Deci, E. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development and well-being. *American Psychologist, 55*(1), 68-78. https://doi.org/10.1037/0003-066X.55.1.68

Ryan, R. M., & Deci, E. L. (2004). Autonomy is no illusion: Self-determination theory and the empirical study of authenticity, awareness, and will. In J. Greenberg, S. L. Koole, & T. Pyszczynski (Eds.), *Handbook of experimental existential psychology* (pp. 449-479). Guilford Press.

Ryan, R. M., & Deci, E. L. (2017). Self-determination theory: An introduction and overview. In R. M. Ryan & E. L. Deci (Eds.), *Basic psychological needs in motivation, development, and wellness* (pp. 3-25). Guilford Press. https://doi.org/10.1521/978.14625/28806

Salsali, M., Tajvidi, M., & Ghiyasvandian, S. (2013). Critical thinking dispositions of nursing students in Asian and non-Asian countries: A literature review. *Global Journal of Health Science, 5*(6), 172-178. https://doi.org/10.5539/gjhs.v5n6p172

Schermelleh-Engel, K., Moosbrugger, H., & Müller, H. (2003). Evaluating the fit of structural equation models: Tests of significance and descriptive goodness-of-fit measures. *Methods of Psychological Research Online, 8*(2), 23-74.

Schleicher, A. (2020). *Teaching and learning international survey - TALIS 2018: Insights and interpretations.* OECD Publishing. https://www.oecd.org/education/talis/TALIS2018_insights _and _interpretations.pdf

Sevari, K., & Farzadi, F. (2022). Examining the model of relationship between professors' support for autonomy and parental psychological control with critical thinking through mediation of academic stress among female students. *Journal of Woman and Culture, 13*(52), 1-15.

Sheldon, K. M., & Krieger, L. S. (2007). Understanding the negative effects of legal education on law students: A longitudinal test of self-determination theory. *Personality and Social Psychology Bulletin, 33*(6), 883–897.https://doi.org/10.1177/0146167207301014

Shen, B., McCaughtry, N., Martin, J., & Fahlman, M. (2009). Effects of teacher autonomy support and students' autonomous motivation on learning in physical education. *Research Quarterly for Exercise and Sport*, *80*(1), 44-53. https://doi.org/10.1080/02701367.2009.10599528

Skaalvik, E. M., & Skaalvik, S. (2014). Teacher self-efficacy and perceived autonomy: Relations with teacher engagement, job satisfaction, and emotional exhaustion. *Psychological Reports*, *114*(1), 68-77. https://doi.org/10.2466/14.02.PR0.114k14w0

Sökmen, Y., & Kılıç, D. (2019). The relationship between primary school teachers' self-efficacy, autonomy, job satisfaction, teacher engagement and burnout: A model development study. *International Journal of Research in Education and Science (IJRES), 5*(2), 709-721.

Stajkovic, A. D., & Luthans, F. (1998). Social cognitive theory and self-efficacy: Going beyond traditional motivational and behavioral approaches. *Organizational Dynamics, 26*(4), 62-74. https://doi.org/10.1016/S0090-2616(98)90006-7

Strong, L. E., & Yoshida, R. K. (2014). Teachers' autonomy in today's educational climate: Current perceptions from an acceptable instrument. *Educational Studies*, *50*(2), 123-145. https://doi.org/10.1080/00131946.2014.880922

Su, Y. L., & Reeve, J. (2011). A meta-analysis of the effectiveness of intervention programs designed to support autonomy. *Educational Psychology Review, 23*(1), 159-188. https://doi.org/10.1007/s10648-010-9142-7

Tschannen-Moran, M., & Hoy, A. W. (2007). The differential antecedents of self -efficacy beliefs of novice and experienced teachers. *Teaching and Teacher Education, 23*(6), 944-956. https://doi.org/10.1016/j.tate.2006.05.003

Tunçeli, H. İ., Yorulmaz, A., & Şahan Aktan, B. (2022). Teacher candidates´ critical thinking and learning autonomy: The mediating role of self-regulation. *Journal for Educators, Teachers and Trainers, 13*(3), 100-110. https://doi.org/ 10.47750/jett.2022.13.03.011

Yang, D., Chen, P., Wang, H., Wang, K., & Huang, R. (2022). Teachers' autonomy support and student engagement: A systematic literature review of longitudinal studies. *Frontiers in Psychology*, (13), 925955. https://doi.org/10.3389/fpsyg.2022.925955

Yazıcı, A. Ş. (2016). The relationship between the teacher autonomy and learner autonomy support behaviors. *Journal of Educational Sciences Research, 6*(2), 1-23. https://doi.org/10.12973/jesr.2016.62.1

Yukselir, C., & Ozer, O. (2022). Investigating the interplay between English language teachers' autonomy, well-being and efficacy. *Issues in Educational Research*, *32*(4), 1643-1657.

Zangenehvandi, M., Farahian, M., & Gholami, H. (2014). On the relationship between EFL teachers' critical thinking and self-efficacy. *Modern Journal of Language Teaching Methods*, *4*(2), 286-298.

Zhang, H., & Lambert, V. (2008). Critical thinking dispositions and learning styles of baccalaureate nursing students from China. *Nursing & Health Sciences, 10*(3), 175-181. https://doi.org/10.1111/j.1442-2018.2008.00393.x

Education and Science

Original Article

Vol 51 (2026) No 226 441-460

The relationship between teacher agency and teacher professional learning activities: the mediating role of teacher self-efficacy

Demet Hacımustafaoğlu [1], Elanur Ersan Albayrak [2], Berrenur Dinç [3], Murat Özdemir [4]

Abstract

Teacher professional learning activities have positive consequences for school outcomes. This study investigated the extent to which teacher agency is correlated with teacher professional learning activities using data from 582 teachers working in lower secondary schools in Ankara, Türkiye. Ankara, as the capital of Türkiye and a major educational hub, was chosen for its strong representational capacity. Using structural equation modeling (SEM), the relationship between teacher agency and teachers' professional learning activities, as well as the mediating role of teacher self-efficacy in this relationship, was tested. The results supported the mediation model, showing that teacher agency was related to teacher professional learning activities both directly and indirectly through teacher self-efficacy. These findings provide an in-depth understanding of how teacher agency is linked to teacher professional learning activities in the context of public schools in Ankara. We also offer practical recommendations for policymakers and school settings to promote teacher professional learning activities and enhance school outcomes.

Keywords

Teacher professional learning activities

Teacher agency

Teacher self-efficacy

Mediation test

Lower secondary school teacher

Article Info

Received: 10.06.2024

Accepted: 11.11.2025

Published Online: 04.30.2026

DOI: 10.15390/ES.2026.2468

Introduction

Teachers' participation in professional learning activities are critical for keeping their knowledge fresh and improving their teaching methods. Teacher education plays a critical role in pre-service as well as in-service education. The continuity of teacher training is important for the quality of education and this is a major concern not only in European countries but worldwide. Previous research points to the vital importance of teacher quality on educational outcomes (Helgevold, 2016; Pang, 2011). Research shows that effective teachers influence not only students' academic achievement but also their social lives (World Bank, 2022). Supporting students academically and socially and contributing to their overall development are the expectations educational stakeholders have from teachers. Meeting these expectations depend on teachers' qualifications and competencies (Lee & Lee, 2020). Therefore, the degree of importance that teachers attribute to their own professional development has implications not only for their personal growth but also for the academic and social development of their students. While existing literature has extensively explored the relationship between teachers' professional learning activities and educational outcomes (Akiba & Liang, 2016; Geleta & Raju, 2023; Rani et al., 2023), there

[1] Hacettepe University, Graduate School of Educational Sciences, Ankara, Türkiye, demethmo@gmail.com
[2] Hacettepe University, Graduate School of Educational Sciences, Ankara, Türkiye, elaerrsan@gmail.com
[3] Hacettepe University, Graduate School of Educational Sciences, Ankara, Türkiye, dincberre@hotmail.com
[4] Hacettepe University, Faculty of Education, Department of Educational Sciences, Ankara, Türkiye, mrtozdem@gmail.com

remain significant gaps in understanding how such activities are shaped in real classroom settings—particularly in relation to teacher agency and self-efficacy. In addition, it has been determined that the quality of teachers has a significant impact on both student achievement and the effectiveness of education (Rockoff, 2004). Likewise, improved pedagogical practices and school-level progress have been positively correlated with teacher learning, defined as a continuous reflective and collaborative professional development process (Avalos, 2011). Initiatives aimed at teachers' professional development have also been crucial in promoting changes in education, increasing teachers' motivation, and ultimately improving student achievement (Desimone & Garet, 2015). Given that professional learning activities serve as a vital mechanism for improving instructional quality and the broader educational experience, it is imperative for future research to investigate the complex interplay among teacher learning, agency, and self-efficacy. To this end, scholarly attention should focus on identifying multilevel determinants—at the teacher, classroom, and school levels—that shape meaningful teacher engagement in professional learning.

A review of previous research on teacher professional learning activities reveals a focus on various aspects such as how teachers learn in schools (Meirink et al., 2009), processes contributing to teacher learning (Voogt et al., 2011), factors influencing teachers' participation in professional learning activities (Kwakman, 2003), current-century teacher education practices (Lieberman & Pointer Mace, 2010), relationships between teachers' professional learning activities and classroom culture (Supovitz & Turner, 2000), and the effects of organizational conditions and leadership practices on teachers' professional learning (Geijsel et al., 2009). These studies highlight an emphasis on examining factors influencing teacher professional learning activities activities and their relationship with other school factors. Despite most teachers striving to maintain high standards in educational activities, they often struggle to reflect this in their professional learning activities (Garet et al., 2001). One reason for teachers' difficulty in translating their support for high educational standards into professional learning activities are believed to be teacher agency. Literature on teacher agency has explored relationships between teacher identity (Choi, 2022), professional development (Lai et al., 2016), and school leadership and agency (Al-Mahdy et al., 2023; Al-Mahdy et al., 2024; Hilal et al., 2022; Polatcan, 2021). However, to the best of our knowledge, no empirical study to date has directly examined the relationship between teacher agency and teacher professional learning activities, nor has it explored the mediating role of teacher self-efficacy in this relationship. Although various studies have examined teacher professional learning in relation to either agency or self-efficacy independently, there is a notable lack of research integrating all three constructs within a single mediational framework.

Although the concepts of teacher agency and self-efficacy have been extensively discussed in previous research (e.g., Buxton et al., 2015; Eteläpelto et al., 2014; Imants & Van der Wal, 2020), empirical studies that integrate both constructs within a single mediational model to explain teachers' engagement in professional learning activities remain scarce. Most existing work has either examined these constructs separately or approached them from a conceptual or theoretical angle. To address this gap, our study empirically tests a model in which teacher agency influences professional learning through the mediating role of self-efficacy. This approach provides a more comprehensive understanding of how internal psychological mechanisms shape teacher learning, particularly within the centralized education system of Türkiye.

The relationship between teacher agency and teacher professional learning activities are likely to be shaped by several mediating variables rather than operating through a direct pathway. Indeed, prior studies have indicated that the influence of teacher agency on various teacher-related outcomes is often exerted through mediational mechanisms (Hilal et al., 2022; Mifsud & Vella, 2018), one of which is teacher self-efficacy (Polatcan et al., 2023). However, there remains a notable gap in the literature regarding how teacher self-efficacy functions within the relationship between teacher agency and teacher professional learning activities, specifically, whether it serves a mediating role. Existing research has examined teacher self-efficacy in relation to teacher enthusiasm (Burić & Moe, 2020; Michos et al.,

2022), classroom management (Chao et al., 2017; Hettinger et al., 2021), and instructional practices (Woodcock et al., 2022). In addition, its mediating role has been explored in the context of leadership and agency (Özdemir et al., 2023; Polatcan et al., 2023), student achievement (Kılınç et al., 2023a), teacher outcomes (Ahn & Bowers, 2024), school climate and teacher stress (Hu et al., 2019), and motivation and job satisfaction (Chang & Sung, 2024). However, there has been no study to examine whether teacher self-efficacy mediates the relationship between teacher agency and professional learning activities. It is particularly important to investigate these concepts in the K-12 context in Türkiye. Türkiye has undergone significant educational transformations in recent years, including curriculum restructuring, decentralization efforts, and reforms aimed at increasing teacher autonomy. Furthermore, Turkish teachers work within a unique sociocultural and policy framework that shapes their agency feelings and professional learning behaviors. Theoretical insights and practical implications for similar centralized education systems can be gained by understanding how these dynamics emerged in Türkiye. Despite these developments, empirical studies examining how teacher agency, self-efficacy, and professional learning interact in this context remain limited. In response to these gaps, the present study aims to explore the relationship between teacher agency and professional learning activities and to determine the mediating role of teacher self-efficacy within this relationship. The following research questions will be addressed:

1. To what extent is there a statistically significant relationship between teacher agency and professional learning activities?
2. To what extent does teacher self-efficacy mediate the relationship between teacher agency and professional learning activities?

By addressing this underexplored intersection of teacher agency, self-efficacy, and teacher professional learning activities, the present study offers several meaningful contributions to both theory and educational practice. First of all, it is assumed that it will positively affect the teacher's role and increase student achievement by contributing to teachers' classroom teaching activities. In a world where student achievement is so important, there is no study in the literature that examines the relationship between teacher professional learning activities and teacher agency, along with the mediating effect of teacher self-efficacy in this relationship. This study is expected to make multifaceted contributions to both the literature and practice. It will raise awareness among administrators, school principals, and teachers about the causes of teacher professional learning activities and will enable them to act more consciously during implementation. This study will fill a gap in the literature with its results. Additionally, the study is expected to contribute to policymakers by examining the factors affecting teachers' professional learning activities and aiding in the creation of educational policy on this issue. Furthermore, investigating these relationships in the Turkish K–12 context is especially relevant given ongoing curricular reforms, the growing emphasis on teacher autonomy, and shifts in national policy that position professional learning as central to educational quality. Despite these developments, empirical research examining how teacher agency, self-efficacy, and professional learning intersect within Türkiye's centralized education system remains scarce. Understanding these dynamics may yield valuable insights for policymakers and practitioners seeking to strengthen teacher development and improve student outcomes.

Conceptual Framework

This study, which investigated the mediating effect of teacher self-efficacy in the relationship between teacher agency and professional learning activities, is grounded in the theoretical framework of self-efficacy, a concept developed within Albert Bandura's (1977) Social Learning Theory, specifically the concept of self-efficacy. Self-efficacy is defined as "the belief in one's capabilities to organize and execute the courses of action required to manage prospective situations" (Bandura, 1995). Bandura (1997) argues that individual abilities vary, leading to the consideration of self-efficacy as a series of differentiated self-beliefs across various domains of functioning. There are four primary sources of self-

efficacy: physiological and emotional states, social persuasion, vicarious experiences, and mastery experiences (Marschall & Watson, 2022). Teacher self-efficacy in particular, is associated with the degree to which teachers feel competent to influence instructional processes and student outcomes (Henson, 2001). In this study, we propose that teacher self-efficacy influences teacher professional learning activities based on these relationships.

Bandura (1986) further emphasizes that human agency emerges from the interaction between individual capacities and environmental conditions. Within this framework, teacher agency refers to the capacity of teachers to act purposefully and autonomously in shaping their professional roles. Bandura's (2001) agentic perspective highlights a reciprocal relationship between self-efficacy and agency: individuals' efficacy beliefs enable agentic action, while agentic experiences in turn shape and reinforce efficacy beliefs. In this context, we posit that teacher agency plays a formative role in the development of self-efficacy, particularly in relation to professional growth and learning. While previous research has reported a directional path from teacher self-efficacy to teacher agency (Polatcan et al., 2023), our study proposes the reverse—namely, that teacher agency is correlated with teacher self-efficacy**.** This theoretical stance is consistent with Bandura's conceptualization, which underlines the active, intentional, and reflective dimensions of agency in shaping individuals' beliefs about their own effectiveness. Accordingly, this study conceptualizes the relationship between teacher agency and self-efficacy as dynamic and mutually reinforcing, while positioning agency as the antecedent variable that contributes to the development of efficacy beliefs within the scope of teachers' professional learning.

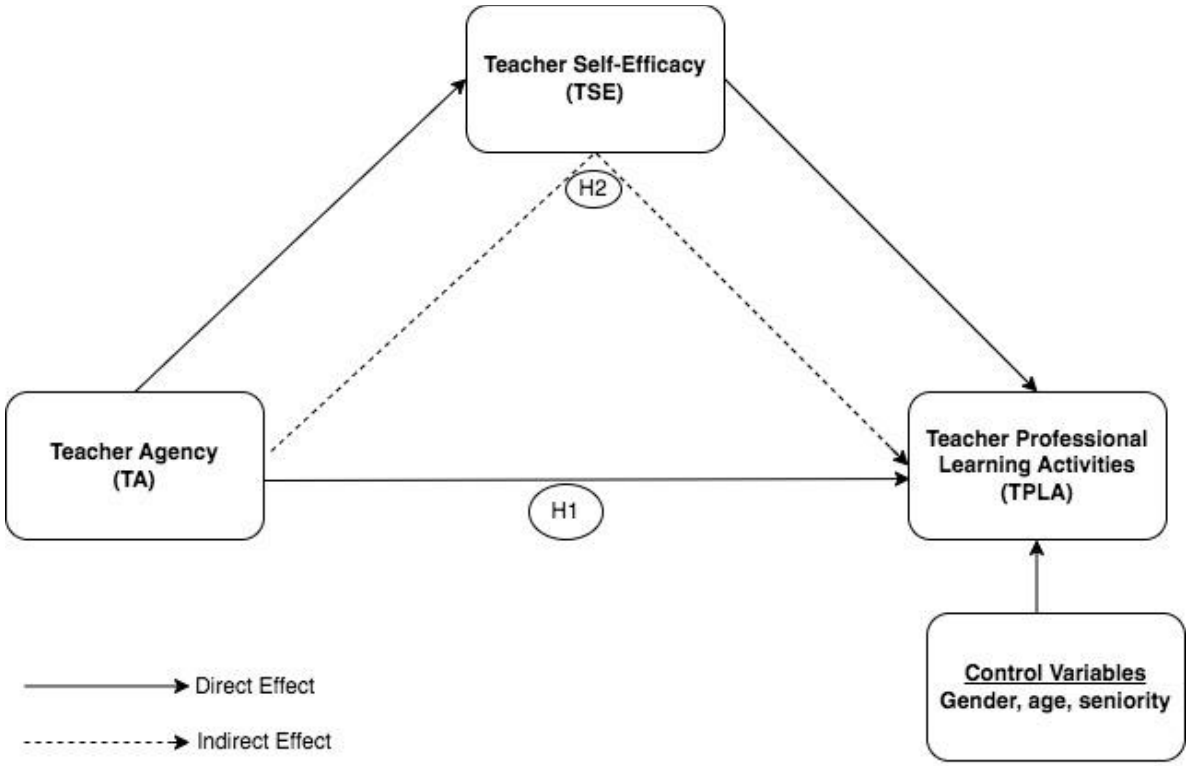


Figure 1. The proposed mediation model linking teacher agency and teacher professional learning activities

Teacher agency and teacher professional learning activities

Recent studies in the field of education have underscored the critical importance of teacher agency (Chung, 2023; Hendawy et al., 2024; Li & Ruppar, 2021). However, the lack of understanding and a clear definition of how agency manifests makes it challenging to differentiate between teachers who demonstrate agency and those who do not (Aspbury-Miyanishi, 2022). Agency can be defined as the capacity to control one's professional practices, work environments, and professional identity (Shavard, 2022). Teacher agency is defined as the ability of teachers to make choices inside and outside the classroom, to take informed action, and to show willingness in their professional development journey (Toom et al., 2015). Agency empowers teachers to initiate their professional activities by granting them the freedom to act (Molla & Nolan, 2020). Teaching activities conducted within a conventional framework are intertwined with various aspects of professional development, rendering teacher agency multidimensional (Tao & Gao, 2017).

Teacher agency refers to teachers' tendency to take initiative, use their internal motivation, and make deliberate and planned efforts to contribute to the school's progress (Liu et al., 2016). In our study, we use Liu, Hallinger, and Feng (2016) teacher agency model to investigate teacher agency. This perspective examines teacher agency in four dimensions: learning efficacy, teaching efficacy, optimism, and constructive engagement. Learning efficacy refers to teachers' ability to improve their learning activities in every situation, including perseverance, motivation, and abilities (Emirbayer & Mische, 1998). Teaching efficacy is described as teachers' willingness to participate in their students' learning process (Frost, 2006). Optimism is defined as positively evaluating the future and retaining positive beliefs about the causes of events (Seligman, 2018). Constructive engagement refers to teachers' level of participation in professional learning processes in a planned, active, and goal-oriented manner (Liu et al., 2016).

Previous studies have emphasized the significant effects of teacher agency. Teacher agency affects teachers' perceptions of the function of education and their approach to students (Biesta et al., 2015). The observations show that teachers who exhibit agency act autonomously in adhering to the curriculum (Aşçı & Yıldırım, 2020). Furthermore, it has been found that the level of teacher agency affects teachers' participation in professional learning activities (Liu et al., 2016). According to the results of literature studies, teacher agency is expected to be related to teachers' professional learning activities.

Teacher professional learning activities are a process that continues throughout their careers and progresses in parallel with their beliefs about learning (Meirink et al., 2010). This process should involve a transition to strategies that are more effective than traditional ones (Korthagen, 2017). Teachers' learning should not only involve the transfer of knowledge, but also the creation of appropriate learning environments for teachers, the ownership of learning outcomes, and the provision of support (Kwakman, 2003). The participation of teachers in professional learning activities are crucial for improving the quality of education, enhancing the teaching profession's quality, and addressing deficiencies in students' learning processes (Akiba, 2015).

Teacher professional learning activities refers to a process in which teachers acquire new knowledge, skills, and values to enhance their educational services (Geijsel et al., 2001). In our study, teacher professional learning activities are examined through the framework developed by Geijsel et al. (2009). Within this framework, professional learning activities are categorized into three dimensions: staying current, experience and reflection, and changing teaching practices. Staying current evaluates the extent to which teachers engage in various activities to develop themselves professionally, emphasizing the importance of staying informed about recentdevelopments in the field. The dimension of experience and reflection encompasses items related to teachers taking action to impart knowledge to students and reflecting on their practices, ultimately resulting in higher quality teaching practices (Geijsel et al., 2001, 2009). The dimension of changing teaching practices focuses on whether and to what extent teachers adapt their teaching methods within a certain period to meet the diverse learning needs of students, including motivating students, diversifying teaching strategies used in the classroom, and considering students' cultural backgrounds and emotional states (Geijsel et al., 2009; Kwakman, 2003).

One significant reason for the relationship between teacher professional learning activities and teacher agency is teacher agency itself. It is expected that for a teacher to demonstrate agency, they should have engaged in personal development and supported their professional learning activities. For instance, a study by Brodie (2021) found that agency facilitates teachers' participation in professional learning communities. The positive impact of teacher agency on professional learning has also been clearly demonstrated (Polatcan, 2021; Wild et al., 2018). This association between teacher agency and teacher professional learning activities are expected to promote active participation in learning activities and foster a dynamic learning environment. Since Bandura (1995, 1997, 2001) defines individuals not merely as passive entities that react to environmental conditions, but as active agents capable of influencing their own behavior, environment, and living conditions. In our study, it is hypothesized that there is a relationship between teacher agency and teacher professional learning activities (H1).

Teacher Self-Efficacy as a Mediator

Individuals' beliefs in their own efficacy, which determine how they interpret situations and events, are the fundamental determinants of their behavior (Zee & Koomen, 2016). Individuals with high self-efficacy are those who take responsibility for their actions and can control them (Wang et al., 2015). Teachers' abilities and self-efficacy influence the structuring of the classroom environment and students' intellectual maturity (Bandura, 1997). Teacher self-efficacy is crucial because it affects students' motivation, persistence in new or challenging situations, and ability to achieve goals (Haworth et al., 2015). Teachers with self-efficacy express themselves freely and create an environment conducive to effective teaching by planning classroom activities (Huang et al., 2020).

The concept of teacher self-efficacy has been frequently studied in the Turkish context. These studies have explored variables such as teacher self-efficacy in relation to classroom management skills (Bayraktar & Çelik, 2021; Çelik, 2019; Demirtaş & Kahveci, 2010; Sak, 2015), leadership (Demir, 2018; Erdogan, 2023; Sağır & Tutkun, 2017), and teacher competencies and professional development (Bozgün & Can, 2021; Keskin & Aktay, 2021; Şeref & Çinpolat, 2021). Findings from these studies suggest that teacher self-efficacy is not merely an individual belief, but is also closely associated with the teaching process and professional attitudes.

Teacher self-efficacy refers to the belief in a teacher's ability to exhibit patience and success in any situation (Tschannen-Moran et al., 1998). In our study, teacher self-efficacy is approached from the perspective of Tschannen-Moran and Hoy (2001). In this perspective, teacher self-efficacy is examined across three dimensions: self-efficacy for student engagement, self-efficacy for using instructional strategies, and self-efficacy for classroom management. These dimensions reflect the richness of teachers' work lives and the requirements of good teaching (Tschannen-Moran & Hoy, 2001).

Although the concepts of teacher agency, self-efficacy, and professional learning activities have different theoretical foundations, some of their characteristics may appear to be interconnected or overlapping in practice. For instance, within the teacher agency framework, teaching activity reflects teachers' deliberate and proactive efforts to influence students' learning outcomes (Liu et al., 2016). However, self-efficacy in using teaching strategies refers to teachers' belief that they can effectively apply specific pedagogical techniques (Tschannen-Moran & Hoy, 2001). While both focus on instructional competence, the former emphasizes goal-oriented actions and conscious behaviors (Bandura, 2001). On the other hand, the latter is based on the self-perception of ability and control over teaching practices (Bandura, 1997). Likewise, the optimism dimension of the teacher agency relates to teachers' positive outlook and future expectations (Seligman, 2018). Since both encompass teachers' sustained motivation and forward-looking engagement, they may be conceptually compatible with the reflective practice component of professional learning activities (Geijsel et al., 2009).

Teacher self-efficacy belief is related to how competent and confident a teacher feels and a teacher with high self-efficacy beliefs believes that they can successfully fulfill their duties; this belief positively affects the teacher's performance (Lemon & Garvis, 2016). In this study, we focus on the mediating effect of teacher self-efficacy in the relationship between teacher agency and teacher professional learning activities. It is believed that if teachers exhibit agency, their self-efficacy beliefs will be affected, which in turn will impact teacher professional learning activities. Therefore, our study hypothesizes that teacher self-efficacy mediates the relationship between teacher agency and teacher professional learning activities (H2).

Method

Reserch Design

This study was designed using a survey research model. A survey model refers to research procedures conducted on an entire population or a sample selected from it, with the aim of reaching generalizable conclusions in a population composed of a large number of elements (Karasar, 2020). Among survey designs, model testing was adopted in this study to examine the hypothesized relationships between variables. In this study, structural equation modeling (SEM) was used to test the hypothesized relationships among teacher agency, teacher self-efficacy, and professional learning activities. SEM refers to a family of related statistical techniques used to examine complex relationships among observed and latent variables (Kline, 2023). This approach allows for the examination of both direct and indirect effects within a multivariate framework, enhancing the understanding of complex educational processes.

Procedure and participants

To ensure the ethical appropriateness of the study, an application was submitted to the Ethics Commission of the Senate of Hacettepe University. The Commission reviewed the study and granted ethical approval during its meeting held on March 6, 2023. Subsequently, official permission to conduct the research in public middle schools was obtained from the Ankara Provincial Directorate of National Education, which oversees the study's target population. Following the approval processes, the researchers visited the schools and provided participants with informed consent forms. Data were then collected face-to-face from those who voluntarily agreed to participate.

This study was situated within lower secondary schools, which constitute a critical stage of compulsory education characterized by structured curricular expectations and observable student academic performance. Moreover, teachers at this level often demonstrate an increased need for professional learning activities, making this context particularly suitable for exploring the variables under investigation. Data were collected from public lower secondary schools located in Ankara, the capital of Türkiye. Ankara was deliberately selected as the research site due to its central administrative role, demographic diversity, and strong representational capacity within the national education system. The choice of the lower secondary level and this specific location aligns with the study's goal of producing findings that are both contextually relevant and broadly applicable. Details regarding the population, sampling, and demographic characteristics of the participants are presented in Table 1.

Table 1. Population, Sampling and Participants (*N*=582)

District	Number of Schools	Number of Teachers	Percentage (%)	Sample Size	Participants Reached
Çankaya	67	2212	12.87	48	75
Altındağ	50	1661	9.66	37	55
Etimesgut	40	2033	11.83	45	62
Gölbaşı	25	514	2.99	11	35
Keçiören	69	3210	18.67	70	97
Mamak	68	2320	13.50	51	93
Pursaklar	18	745	4.33	16	29
Sincan	51	2213	12.87	48	62
Yenimahalle	57	2283	13.28	50	74
Total	445	17.191	100%	376	582

Table 1. Continued

	N	%
Gender		
Male	135	23.2
Female	447	76.8
Seniority (year)		
<=5	29	5
6-10	66	11.3
11-15	126	21.6
16-20	150	25.8
21=>	210	36.1
Age		
<=30	37	6.4
31-40	161	27.7
41-50	258	44.3
51=>	125	21.5
Education		
Bachelor's degree	479	82.3
Graduate	90	15.5

As shown in Table 1, the study population comprised 17,191 teachers working in 445 public lower secondary schools across nine central districts of Ankara during the 2023–2024 academic year. The sample was determined using a proportionate stratified sampling method, with the sample size from each district reflecting its relative share of the overall teacher population. This process resulted in a total sample size of 376 teachers. To ensure strong representativeness, districts with the largest teacher populations were prioritized. Accordingly, nine out of the ten districts with the highest number of teachers in Ankara were included in the sample. One district, Polatlı, was purposefully excluded despite its large teacher population due to its geographic remoteness, which posed substantial logistical challenges for in-person data collection. All other selected districts were located within or in close proximity to the urban center, thereby facilitating a more feasible and consistent data collection process. In total, data collection was carried out in 35 randomly selected public lower secondary schools across the sampled districts. In the implementation phase of the study, 600 data collection tools were randomly distributed to teachers in 35 schools. The data were collected by the researchers by visiting the schools, using face-to-face method and with the voluntary participation of the teachers. As a result, data were returned from 582 participants and data were obtained.

As shown in the table, 76.8% of the participants were female. This proportion is consistent with broader national and regional trends. According to national statistics for the 2023–2024 academic year, 58.29% of teachers working in public schools across Türkiye were female, while the proportion rose to 72.37% in Ankara (Ministry of National Education [MoNE], 2024). These figures suggest that the gender distribution in the sample aligns with the actual teacher population, particularly in metropolitan areas such as Ankara and Istanbul, where the teaching profession is predominantly occupied by women. In addition nearly half of the participants were in the middle age group. Considering their educational background, bachelor's degree constitutes a significant majority of the participants. Finally, the professional experience of the participants shows a balanced distribution

Measures

Teacher Agency: The Teacher Agency Scale (TAS), developed by Liu et al. (2016) and adapted to the Turkish culture by Bellibaş et al. (2019), is used to assess teacher agency. TAS is a four-dimensional scale consisting of 24 items. The dimensions of TAS are (a) learning efficacy, (b) teaching efficacy, (c) optimism, and (d) constructive participation. TAS is a 5-point Likert-type scale with scores varying between strongly *disagree* (1) to *strongly agree* (5). A sample item of the TAS is "*Even when I feel bad, I can*

actively participate in professional learning activities." CFA was conducted to test the structural validity of TAS. The results indicated a good fit for the four-factor structure (RMSEA=0.05, CFI=0.91, TLI=0.90, SRMR=.07, χ2/sd =2,38). According to Hu and Bentler (1999), these fit indices fall within the acceptable thresholds (CFI and TLI ≥ 0.90, RMSEA ≤ 0.08, SRMR ≤ 0.08), suggesting a satisfactory model fit. The alpha coefficient for this scale is 0.90. Given that the TAS comprises four distinct dimensions, composite reliability was calculated, yielding a coefficient of 0.933.

Teacher Self-Efficacy: The Teacher Self-Efficacy Scale (TSES), developed by Tschannen-Moran and Hoy (2001) and adapted to the Turkish culture by Karaoğlu (2019), is used to assess teacher self-efficacy. TSES is a three-dimensional scale consisting of 12 items. The dimensions of TSES are (a) self-efficacy for student engagement, (b) self-efficacy for using instructional strategies, and (c) self-efficacy for classroom management. TSES is a 9-point Likert-type scale with scores varying *between nothing* (1) and *a great deal* (9). A sample item of the TSES is "*How much effort can you put into convincing your students to succeed in school?*" CFA was conducted to test the structural validity of TSES. The results indicated a good fit for the four-factor structure (RMSEA=0.078, CFI=0.92, TLI=0.90, SRMR=.05, χ2/sd = 4,46). According to Hu and Bentler (1999), these fit indices fall within the acceptable thresholds (CFI and TLI ≥ 0.90, RMSEA ≤ 0.08, SRMR ≤ 0.08), suggesting a satisfactory model fit. The alpha coefficient for this scale is 0.88. Given that the TSES comprises three distinct dimensions, composite reliability was calculated, yielding a coefficient of 0.909.

Teacher professional learning activities: The Teacher Professional Learning Activities Scale (TPLAS), developed by Geijsel et al. (2009) and adapted to the Turkish culture by Polatcan (2020), is used to measure teachers' professional learning activities. TPLAS is a three-dimensional scale consisting of 17 items. The dimensions of TPLAS are (a) staying current, (b) experience and reflection, and (c) changing instruction. TPLAS is a 4-point Likert-type scale with scores varying between *never* (1) and *always* (4). A sample item of the TPLAS is "*I participate in in-service training activities even if they are not mandatory*" CFA was conducted to test the structural validity of TPLAS. The results indicated a good fit for the four-factor structure (RMSEA=0.054, CFI=0.91, TLI=0.90, SRMR=.04, χ2/sd =2,66). According to Hu and Bentler (1999), these fit indices fall within the acceptable thresholds (CFI and TLI ≥ 0.90, RMSEA ≤ 0.08, SRMR ≤ 0.08), suggesting a satisfactory model fit. The alpha coefficient for this scale is 0.86. Given that the TPLAS comprises three distinct dimensions, composite reliability was calculated, yielding a coefficient of 0.898.

Control Variables: Previous literature suggests a connection between teachers' gender, seniority, and their professional practices (Kılınç et al., 2023b). In our study, we have chosen to include gender, seniority, and age as control variables.

Analytic strategies

The 600 data collection tools gathered were transferred to the SPSS 26 program. Out of the 600 tools, missing and outliers were identified, leading to the elimination of 18 data collection tools. Consequently, analyses were continued on the remaining 582 data. Although the scales employed in this study were originally developed as multidimensional, composite scores were used in the analyses. This decision was theoretically grounded in Bandura's (1977) Social Learning Theory, which conceptualizes constructs such as teacher agency and self-efficacy as overarching latent traits influencing multiple related behaviors. Consistent with the guidelines provided by Bagozzi and Edwards (1998), the use of composite scores is appropriate when the subdimensions demonstrate conceptual unity and empirical consistency. In this study, the subdimensions within each scale exhibited high internal consistency, with Composite Reliability (CR) and Cronbach's alpha (α) values exceeding recommended thresholds. These results, along with the theoretical focus on the overall constructs rather than their specific dimensions, justify the use of composite scores for subsequent analyses. The data analysis process began with the analysis of descriptive statistics. Following this, Confirmatory Factor Analysis (CFA) was applied to the proposed model using Mplus 8.3 (Muthén &

Muthén, 2017) to provide structural discrimination between variables. Maximum likelihood (MLR) was used as the estimator in all analyses. The fit values of the research were examined. Since the research data were obtained from a single source—teachers—certain measures were taken to reduce common method bias (CMB). To manage CMB, we controlled for the effects of a single unmeasured latent method factor, as suggested by Lindell and Whitney (2001) and Podsakoff et al. (2012). This approach involved performing the common-method bias test by controlling for the effects of a single unmeasured latent method factor.

Next, we used structural equation modeling (SEM) to test the mediation model for the effects of teacher agency on teacher professional learning activities through teacher self-efficacy. At this stage, model fit was examined using CFI, TLI, and RMSEA, in line with the recommendations of Hu and Bentler (1999). Unidimensional variables were utilized when entering variables into SEM. To assess the linearity assumption, we created scatter plots of the dependent variable against each of the independent variables. We observed that there were no discernible curves or clusters of data points, confirming linearity. To address homoskedasticity, we plotted the residuals against the fitted values. The residuals were evenly distributed around zero without forming a distinct funnel shape, indicating homoskedasticity. Regarding the multicollinearity assumption, we calculated the variance inflation factor (VIF). The VIF values were all below 10, indicating that there was no multicollinearity between the independent variables. Finally, we assessed mediation indices and conditional indirect effects using bootstrap confidence intervals (CIs) based on 5000 bootstrapped samples (Hayes, 2022). Bootstrapping procedures were employed to test the significance of indirect effects, as the distribution of indirect effects is often non-normal (Preacher & Hayes, 2008). Bootstrapping provides a robust, non-parametric method for estimating confidence intervals without relying on the assumption of normality, making it particularly appropriate for mediation analysis in the context of this study.

Results

Descriptive statistics and correlations

Table 2 presents the means, standard deviations and Pearson correlations for the variables of this study. As shown in Table 2, teachers' perceptions of teacher agency (M = 4.02, SD = .42), teacher self-efficacy (M = 7.37, SD = .82), and teacher professional learning activities (M = 3.11, SD = .37) are generally high. These findings suggest that teachers exhibit a common pattern in their levels of agency and engagement in professional learning activities, whereas their perceptions of self-efficacy display relatively greater individual variation. Moreover, there are significant and positive correlations between teacher agency and teacher self-efficacy (r =.45, p< .05), teacher professional learning activities (r =.59, p< .05). Finally, there is a significant and positive correlation between teacher professional learning activities and teacher self-efficacy (r =.48, p< .05). These results show preliminary support for the hypothesis of the study.

Table 2. Descriptive statistics (N=582)

	M	SD	TA	TSE	TPLA
TA	4.02	.42	--		
TSE	7.37	.82	0.56*	--	
TPLA	3.11	.37	0.67*	0.57*	

Abbreviations: M, mean; SD, standard deviation; TA, teacher agency; TSE, teacher self-efficacy; TPLA, teacher professional learning activities; *p<.05.

Result of measurement model and common method bias

The results of the measurement model are presented in Table 3. The three-factor model (TA, TPLA, TSE) demonstrated a superior fit to the data compared to the alternative models ($\chi 2$ =2624.908, df=1279, $\chi 2/df$=2.05 CFI=0.85, TLI=0.84, RMSEA=0.04 and SRMR=0.06). To assess potential common

method bias (CMB), we used the technique of controlling for the effect of a single unmeasured latent method factor (Podsakoff et al., 2012). After controlling for the common latent factor (CLF), there was no substantial improvement in model fit indices ($\chi 2$ = 3722.798, *df*=1446, $\chi 2/df$=2.57, CFI=0.81, TLI=0.79, RMSEA=0.05). Additionally, a chi-square difference test comparing the original measurement model with the CLF-adjusted model revealed a statistically significant difference ($\Delta\chi^2[167] = 1097.89$, $p < .001$). However, given that the inclusion of the CLF did not meaningfully enhance model fit, it can be concluded that common method bias is unlikely to pose a significant threat to the validity of the findings. Furthermore, the empirical distinction among the three constructs was confirmed, supporting their validity as separate latent variables. Furthermore, the comparative analysis of alternative models reinforces the discriminant validity of the three-factor structure. Model 1**,** which posits a unidimensional construct by merging TA, TSE, and TPLA, demonstrated the poorest model fit (χ^2/df = 3.81, CFI = 0.59, TLI = 0.57, RMSEA = 0.07, SRMR = 0.07), indicating substantial misfit**.** Model 2, which combines TA and TSE into a single factor while treating TPLA as distinct, yielded slightly improved yet still unsatisfactory fit indices (χ^2/df = 3.30, CFI = 0.66, TLI = 0.65, RMSEA = 0.06, SRMR = 0.06). In contrast, Model 3**,** representing TA, TSE, and TPLA as three separate latent constructs, exhibited notably superior fit statistics. These results provide robust empirical support for the conceptual distinction among the constructs, affirming their multidimensional nature within the proposed theoretical framework.

Table 3. Model fit results from confirmatory factor analysis.

Variable	$\chi 2$	*df*	$\chi 2/df$	CFI	TLI	RMSEA	SRMR
Model 3: TA;TSE;TPLA	2624.908	1279	2.05	0.85	0.84	0.04	0.06
Model 2: TA+TSE;TPLA	4380.093	1324	3.30	0.66	0.65	0.06	0.06
Model 1: TA+TSE+TPLA	5048.656	1325	3.81	0.59	0.57	0.07	0.07

Abbreviations: TA = teacher agency; TPLA= teacher professional learning activities; TSE= teacher self-efficacy

Hypotheses testing

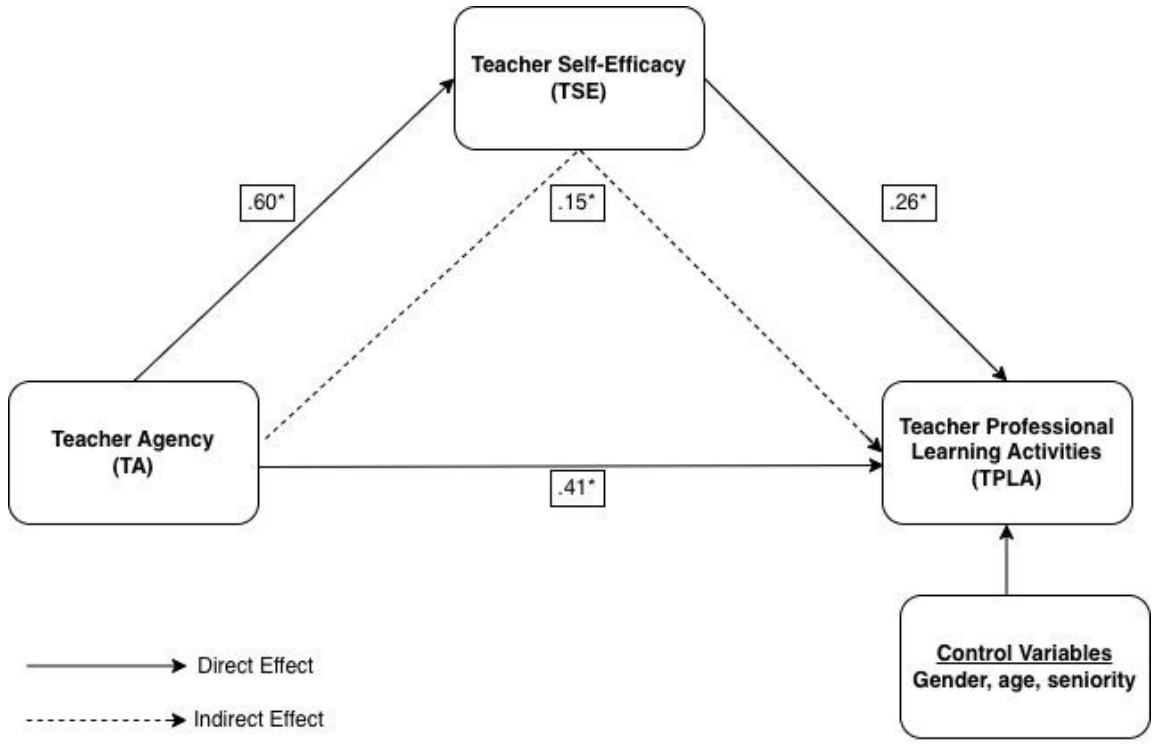
First, we tested whether teacher agency had a directly correlated with teacher professional learning activities. The results showed that teacher agency had a significant correlated with effect on teacher professional learning activities (β = .41, SE = .08, 95% CI [.24, .57], $p < .001$). In addition, teacher self-efficacy was significantly correlated with teacher professional learning activities (β = .26, SE = .05, 95% CI [.16, .35], $p < .001$), while teacher agency also had a s significant direct correlated with on teacher self-efficacy (β = .60, SE = .06, 95% CI [.50, .71], $p < .001$), explaining 36.5% of its variance (R^2 = .365). According to Gignac and Szodorai's (2016) guidelines, this indicates a relatively large effect size. The results confirm Hypothesis 1, indicating a significant relationship between teacher agency and teacher professional learning activities.

To test the mediating role of teacher self-efficacy (Hypothesis 2), we employed a bootstrapping approach with 5000 resamples (Preacher & Hayes, 2008). The analysis revealed that the indirect correlated between teacher agency and teacher professional learning activities through teacher self-efficacy was statistically significant (β = .15, SE = .02, 95% CI [.10, .20], $p < .001$), supporting Hypothesis 2. Approximately 27.6% of the overall relationship was explained by this mediating path, suggesting a medium-sized mediation. Finally, the total standardized relationship between teacher agency on teacher professional learning activities was strong (β = .56, SE = .08, 95% CI [.40, .73], $p < .001$), indicating that teacher agency is linked to professional learning both directly and indirectly. The model explained 37.9% of the variance in teacher professional learning activities (R^2 = .379), which also represents a relatively large effect size (Gignac & Szodorai, 2016). The model demonstrated a good fit to the data (CFI = .99, TLI = .98, RMSEA = .04). The model results are presented in Table 4 and illustrated in Figure 2.

Table 4. Standardised coefficients for testing the direct, mediating effects (N = 582)

	Coefficient		95% confidence interval		
	Estimate	**SE**	**Lower bound**	**Upper bound**	***p***
Direct effects					
TA -- TSE	.60	.06	.50	.71	*
TSE -- TPLA	.26	.05	.16	.35	*
TA -- TPLA	.41	.08	.24	.57	*
Indirect effects					
TA -- TPLA	.15	.02	.10	.20	*
Total effects					
TA -- TPLA	.56	.08	.40	.73	*

TA = teacher agency; TSE= teacher self-efficacy; TPLA= teacher professional learning activities; *p < .05.



*p < .05

Figure 2. Structural model results illustrating the mediating role of teacher self-efficacy in the relationship between teacher agency and teacher professional learning activities.

As illustrated in Figure 2, the structural model visually represents the hypothesized relationships and standardized coefficients, emphasizing the mediating role of teacher self-efficacy in the pathway from teacher agency to professional learning activities. This visual layout complements the statistical results presented in the previous section.

Discussion

This study investigated the mediating role of teacher self-efficacy in the relationship between teacher agency and teacher professional learning activities. Age, gender, and seniority were included as control variables. As expected, there is a significant relationship between teacher agency and teacher professional learning activities. The results also show that teacher self-efficacy mediates the relationship between teacher agency and teacher professional learning activities. Additionally, the findings showed that teacher agency had an indirect correlation with teacher professional learning activities through teacher self-efficacy.

The interconnected roles of teacher agency and teacher self-efficacy in influencing teacher professional learning activities are highlighted in our study, which adds to the existing corpus of literature. The results show that self-efficacy mediates the relationship between professional learning and teacher agency, which is a significant antecedent of professional learning. All of these findings highlight the psychological processes that shape teachers' engagement in learning, particularly within the Turkish educational system where teacher development is a key priority. Teachers with greater agency also exhibit higher levels of self-efficacy, consistent with previous studies (Kim et al., 2018; Min, 2023). This supports the notion that teachers who perceive themselves as capable of initiating change and acting proactively possess greater confidence in their professional abilities. Prior research further indicates that teachers' willingness to participate in professional development is strongly influenced by their self-efficacy (Bray-Clark & Bates, 2003; Geijsel et al., 2009; Huang et al., 2020; Tschannen-Moran & McMaster, 2009), reinforcing the idea that self-efficacy fosters engagement in reflective and innovative learning activities. Additionally, studies have shown that teacher agency directly impacts professional learning (Brodie, 2021; Liu et al., 2016), with teachers who demonstrate higher agency being more likely to pursue opportunities for ongoing learning and growth (Liu et al., 2016).

Our findings further highlight the dynamic interplay between these constructs by demonstrating that teacher self-efficacy mediates the relationship between agency and professional learning activities. This result extends the limited body of research that integrates these three variables into a single mediational framework (Hilal et al., 2022; Mifsud & Vella, 2018) and provides empirical support for Bandura's (2001) Social Cognitive Theory, which posits that human behavior arises from the interaction of behavioral, environmental, and personal factors. A substantial body of research conducted in the Turkish context reinforces these findings, linking teacher self-efficacy to professional competencies (Bozgün & Can, 2021; Keskin & Aktay, 2021; Şeref & Çinpolat, 2021), leadership (Demir, 2018; Erdogan, 2023; Sağır & Tutkun, 2017), and classroom management (Bayraktar & Çelik, 2021; Çelik, 2019; Demirtaş & Kahveci, 2010; Sak, 2015). These studies collectively show that teacher self-efficacy is not merely an individual psychological attribute but is closely tied to teachers' professional behaviors and motivation to engage in development activities. By illustrating that self-efficacy not only influences professional learning but also serves as a psychological mechanism through which teacher agency translates into sustained professional growth, this study underscores its integrative role in promoting active teacher engagement—especially in centralized educational systems like Türkiye's, where opportunities for individual initiative may be constrained.

We theoretically and empirically confirmed the main hypothesis of our study, that teacher agency promotes teacher professional learning through teacher self-efficacy. In Türkiye, teachers often face challenges in taking initiative and exercising agency due to the highly centralized nature of the education system (Özdemir et al., 2023). Teachers with high self-efficacy can take initiative and assert their presence in the classroom. Teachers who lack self-efficacy do not consider the classroom atmosphere and apply the rules from the central system rigidly (Aydın, 2016). Teacher agency is important for the development of teacher professional learning activities. In a centralized country like Türkiye, teachers can only become agents through their own self-efficacy. Therefore, it is important to develop teachers' self-efficacy for teacher professional learning activities. Our study concluded that teacher agency affects teachers' self-efficacy beliefs and, thus, teacher professional learning activities. Therefore, it is essential to foster teachers' self-efficacy in order to support their engagement in professional learning activities. Our findings indicate that teacher agency significantly influences self-efficacy beliefs, which in turn affect participation in teacher professional learning activities.

For teacher professional learning to develop, teachers should be individuals who take initiative and have high self-efficacy beliefs. To achieve teacher professional learning activities, we suggest that teachers should be capable of taking initiative and becoming proactive subjects. From this perspective, teachers' self-efficacy beliefs will begin to increase, allowing them to assert their presence within the school organization. We have also demonstrated the importance of teacher self-efficacy for teacher professional learning activities, emphasizing that teachers should undergo in-service training to nurture their self-efficacy beliefs before they take office. In this way, they will actively participate in teacher professional learning activities as individuals with high self-efficacy once they begin their roles. Finally, to move away from the perception of teachers as passive implementers of centrally mandated decisions—common in Türkiye's centralized education system—teachers should be given greater autonomy to make pedagogical decisions at the local level. Policymakers should support this by empowering teachers with the authority to act independently. Ultimately, such measures have the potential to improve both teacher development and student learning outcomes. This aligns with previous recommendations that empowering teachers locally leads to more sustainable and impactful professional development (Avalos, 2011; Desimone & Garet, 2015).

Educational research is often oriented towards improving student performance. In our study, we focused on the teacher and investigated the direct effects of teacher agency and the indirect effects of teacher self-efficacy on teacher professional learning activities. As a result, our study offers a unique and original contribution to future teacher-related research in the field of education. By grounding our work in Bandura's Social Cognitive Theory and extending it through a mediational model, we offer new insights into the internal mechanisms that foster teacher learning and professional growth. We aimed for our study to set an example for researchers by building on the theoretical foundations of self-efficacy theory.

Limitations

This study has some limitations that may shed light for future researchers. First, in this study, direct and indirect relationships were investigated using a cross-sectional design. Therefore, a cause-and-effect relationship could not be established between the research variables. In the future, researchers could examine the changes in the relationships between variables over time by using a longitudinal research design. Second, the data were collected from only one source, namely teachers. In future research, consulting the views of other members of the school can make an important contribution. Third, only the mediating effect of teacher self-efficacy was investigated. Thus, we suggest future researchers to examine variables such as autonomy, trust, motivation, which may have a moderate impact on this model. Finally, qualitative research designs can be utilized to provide more detailed findings on the nature of the relationship between teacher efficacy and teacher professional learning activities. Fifth, the sample was limited to teachers working in public middle schools in Ankara, Türkiye. Therefore, the generalizability of the findings may be restricted to this specific population. Future research could include participants from different school levels, regions, or school types such as private institutions. Sixth, all data were collected through self-report questionnaires, which may be prone to social desirability bias or common method variance. Using multiple data sources or methods (e.g., interviews, observations) may strengthen the validity of future findings. Finally, although structural equation modeling (SEM) was used, the analysis was limited to linear relationships. Future studies could explore potential non-linear associations or test alternative models to uncover deeper patterns.

References

Ahn, J., & Bowers, A. J. (2024). Do teacher beliefs mediate leadership and teacher behaviors? Testing teacher self-efficacy's mediation role between leadership for learning and teacher outcomes. *Journal of Educational Administration*, *62*(2), 197-222. https://doi.org/10.1108/JEA-12-2022-0227

Akiba, M. (2015). Measuring teachers' professional learning activities in international context. In G. K. Letendre & A. W. Wiseman (Eds.), *Promoting and sustaining a quality teacher workforce* (Vol. 27, pp. 87-110). Emerald Group Publishing Limited. https://doi.org/10.1108/S1479-367920140000027003

Akiba, M., & Liang, G. (2016). Effects of teacher professional learning activities on student achievement growth. *The Journal of Educational Research*, *109*(1), 99-110. https://doi.org/10.1080/00220671.2014.924470

Al-Mahdy, Y. F. H., Al-Harthi, A. S., & Salah El-Din, N. (2023). The effect of leadership support on commitment to change and turnover intention in Omani higher education. *Journal of Further and Higher Education*, *47*(3), 324-337. https://doi.org/10.1080/0309877X.2022.2124367

Al-Mahdy, Y. F. H., Hallinger, P., Emam, M., Hammad, W., Alabri, K. M., & Al-Harthi, K. (2024). Supporting teacher professional learning in Oman: The effects of principal leadership, teacher trust, and teacher agency. *Educational Management Administration & Leadership*, *52*(2), 395-416. https://doi.org/10.1177/17411432211064428

Aspbury-Miyanishi, E. (2022). The affordances beyond what one does: Reconceptualizing teacher agency with Heidegger and Ecological Psychology. *Teaching and Teacher Education*, (113), 103662. https://doi.org/10.1016/j.tate.2022.103662

Aşçı, M., & Yıldırım, R. (2020). Öğretmen failliği ile öğretim programına bağlılık arasındaki ilişki: Yunus Emre ilçesi örneği [Special issue]. *OPUS–Uluslararası Toplum Araştırmaları Dergisi*, (16), 6126-6149. https://doi.org/10.26466/opus.806477

Avalos, B. (2011). Teacher professional development in teaching and teacher education over ten years. *Teaching and Teacher Education*, *27*(1), 10-20. https://doi.org/10.1016/j.tate.2010.08.007

Aydın, Y. (2016). Örgütsel sessizliğin okul yönetiminde kayırmacılık ve öğretmenlerin öz yeterlik algısı ile ilişkisi. *Kuram ve Uygulamada Eğitim Yönetimi Dergisi*, *22*(2), 165-192. https://doi.org/10.14527/kuey.2016.007

Bagozzi, R. P., & Edwards, J. R. (1998). A general approach for representing constructs in organizational research. *Organizational Research Methods*, *1*(1), 45-87. https://doi.org/10.1177/109442819800100104

Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review*, *84*(2), 191. https://doi.org/10.1037/0033-295X.84.2.191

Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory*. PrenticeHall.

Bandura, A. (1995). *Self-efficacy in changing societies.* Cambridge University Press. https://doi.org/10.1017/CBO9780511527692

Bandura, A. (1997). *Self-efficacy: The exercise of control*. W H Freeman/Times Books/ Henry Holt & Company.

Bandura, A. (2001). Social cognitive theory: An agentic perspective. *Annual Review of Psychology*, *52*(1), 1-26. https://doi.org/10.1146/annurev.psych.52.1.1

Bayraktar, H. V., & Çelik, O. (2021). Öğretmenlerin öz yeterlilikleri ile sınıf yönetimi becerileri arasındaki ilişkinin incelenmesi. *İZÜ Eğitim Dergisi*, *3*(6), 98-127. https://doi.org/10.46423/izujed.841318

Bellibaş, M. Ş., Çalışkan, Ö., & Gümüş, S. (2019). Öğretmen Failliği Ölçeği'nin (ÖFÖ) geçerlik ve güvenilirlik çalışması. *Trakya Eğitim Dergisi, 9*(1), 1-11.

Biesta, G., Priestley, M., & Robinson, S. (2015). The role of beliefs in teacher agency. *Teachers and Teaching*, *21*(6), 624-640. https://doi.org/10.1080/13540602.2015.1044325

Bozgün, K., & Can, F. (2021). Öğretmen adaylarının serbest yazma tutumlarında öğretmen öz yeterlik algılarının yordayıcılığı. *RumeliDE Dil ve Edebiyat Araştırmaları Dergisi*, (24), 21-33. https://doi.org/10.29000/rumelide.990059

Bray-Clark, N., & Bates, R. (2003). Self-efficacy beliefs and teacher effectiveness: Implications for professional development. *Professional Educator*, *26*(1), 13-22.

Brodie, K. (2021). Teacher agency in professional learning communities. *Professional Development in Education*, *47*(4), 560-573. https://doi.org/10.1080/19415257.2019.1689523

Burić, I., & Moe, A. (2020). What makes teachers enthusiastic: The interplay of positive affect, self-efficacy and job satisfaction. *Teaching and Teacher Education*, (89), 103008. https://doi.org/10.1016/j.tate.2019.103008

Buxton, C. A., Allexsaht-Snider, M., Kayumova, S., Aghasaleh, R., Choi, Y. J., & Cohen, A. (2015). Teacher agency and professional learning: Rethinking fidelity of implementation as multiplicities of enactment. *Journal of Research in Science Teaching*, *52*(4), 489-502. https://doi.org/10.1002/tea.21223

Chang, T. J., & Sung, Y. T. (2024). Does teacher motivation really matter? Exploring the mediating role of teachers' self-efficacy in the relationship between motivation and job satisfaction. *The Asia-Pacific Education Researcher*, (33), 1315-1325. https://doi.org/10.1007/s40299-023-00803-4

Chao, C. N. G., Sze, W., Chow, E., Forlin, C., & Ho, F. C. (2017). Improving teachers' self-efficacy in applying teaching and learning strategies and classroom management to students with special education needs in Hong Kong. *Teaching and Teacher Education*, (66), 360-369. https://doi.org/10.1016/j.tate.2017.05.004

Choi, J. (2022). The interplay of teacher identities and agency: A case study of two native English-speaking teachers in South Korea. *Teaching and Teacher Education*, (117), 103804. https://doi.org/10.1016/j.tate.2022.103804

Chung, J. (2023). Research-informed teacher education, teacher autonomy and teacher agency: The example of Finland. *London Review of Education*, *21*(1) 1-11. https://doi.org/10.14324/LRE.21.1.13

Çelik, O. (2019). *İlkokul ve ortaokul öğretmenlerinin öz yeterlilikleri ile sınıf yönetimi becerileri arasındaki ilişkinin incelenmesi* (Thesis No. 553475) [Master's thesis, Istanbul Sabahattin Zaim University]. Council of Higher Education National Thesis Center.

Demir, S. (2018). Okul yöneticilerinin motivasyonel dili ile öğretmen öz yeterliği arasındaki ilişki üzerine bir çalışma. *Anemon Muş Alparslan Üniversitesi Sosyal Bilimler Dergisi*, *6*(2), 177-183. https://doi.org/10.18506/anemon.384848

Demirtaş, Z., & Kahveci, G. (2010). Öğrenci algılarına göre 4. sınıf ve 5. sınıf öğretmenlerinin sınıf yönetimi yeterlikleri. *Dicle Üniversitesi Ziya Gökalp Eğitim Fakültesi Dergisi*, (15), 18-29.

Desimone, L. M., & Garet, M. S. (2015). Best practices in teachers' professional development in the United States. *Psychology, Society and Education*, *7*(3), 252-263. https://doi.org/10.25115/psye.v7i3.515

Emirbayer, M., & Mische, A. (1998). What is agency?. *American Journal of Sociology, 103*(4), 962-1023. https://doi.org/10.1086/231294

Erdogan, U. (2023). Öğretmen liderliği ve öz yeterliği arasındaki ilişkinin incelenmesi. *İnönü Üniversitesi Eğitim Bilimleri Enstitüsü Dergisi*, *10*(20), 213-232. https://doi.org/10.29129/inujgse.1362718

Eteläpelto, A., Vähäsantanen, K., Hökkä, P., & Paloniemi, S. (2014). Miten käsitteellistää ammatillista toimijuutta työssä?. *Aikuiskasvatus, 34*(3), 202-214. https://doi.org/10.33336/aik.94100

Frost, D. (2006) The concept of 'agency' in leadership for learning. *Leading & Managing, 12*(2), 19-28.

Garet, M. S., Porter, A. C., Desimone, L., Birman, B. F., & Yoon, K. S. (2001). What makes professional development effective? Results from a national sample of teachers. *American Educational Research Journal*, *38*(4), 915-945. https://doi.org/10.3102/000283120380049

Geijsel, F., Sleegers, P., Van den Berg, R., & Kelchtermans, G. (2001). Conditions fostering the implementation of large-scale innovation programs in schools: Teachers' perspectives. *Educational Administration Quarterly*, *37*(1), 130-166. https://doi.org/10.1177/00131610121969262

Geijsel, F. P., Sleegers, P. J. C., Stoel, R. D., & Krüger, M. L. (2009). The effect of teacher psychological and school organizational and leadership factors on teachers' professional learning in Dutch schools. *Elementary School Journal*, *109*(4), 406-427. https://doi.org/10.1086/593940

Geleta, T. O., & Raju, T. S. (2023). Professional learning activities in the higher education institution of Ethiopia. *Heliyon*, *9*(3), e14119. https://doi.org/10.1016/j.heliyon.2023.e14119

Gignac, G. E., & Szodorai, E. T. (2016). Effect size guidelines for individual differences researchers. *Personality and Individual Differences*, (102), 74-78. https://doi.org/10.1016/j.paid.2016.06.069

Haworth, P., McGee, A., & Maclntyre, L. K. (2015). Building a whole school approach and teacher efficacy with English language learners. *Teachers and Teaching*, *21*(2), 164-177. https://doi.org/10.1080/13540602.2014.928131

Hayes, A. (2022). *Introduction to mediation, moderation and conditional processes analysis: A regression- based approach* (3rd ed.). The Guilford Press.

Helgevold, N. (2016). Teaching as creating space for participation–establishing a learning community in diverse classrooms. *Teachers and Teaching*, *22*(3), 315-328. https://doi.org/10.1080/13540602.2015.1058590

Hendawy Al-Mahdy, Y. F., Hallinger, P., Emam, M., Hammad, W., Alabri, K. M., & Al-Harthi, K. (2024). Supporting teacher professional learning in Oman: The effects of principal leadership, teacher trust, and teacher agency. *Educational Management Administration & Leadership*, *52*(2), 395-416. https://doi.org/10.1177/17411432211064428

Henson, R. K. (2001, January 26). *Teacher self-efficacy: Substantive implications and measurement dilemmas* [Invited keynote address]. Annual meeting of the Educational Research Exchange, Texas A&M University, College Station, TX. https://files.eric.ed.gov/fulltext/ED452208.pdf

Hettinger, K., Lazarides, R., Rubach, C., & Schiefele, U. (2021). Teacher classroom management self-efficacy: Longitudinal relations to perceived teaching behaviors and student enjoyment. *Teaching and Teacher Education*, (103), 103349. https://doi.org/10.1016/j.tate.2021.103349

Hilal, Y. Y., Hammad, W., & Polatcan, M. (2022). Does distributed leadership improve teacher agency? Exploring the mediating effect of teacher reflection. *Educational Management Administration & Leadership*, *52*(6), 1369-1387. https://doi.org/10.1177/17411432221134931

Hu, B. Y., Li, Y., Wang, C., Reynolds, B. L., & Wang, S. (2019). The relation between school climate and preschool teacher stress: The mediating role of teachers' self-efficacy. *Journal of Educational Administration*, *57*(6), 748-767. https://doi.org/10.1108/JEA-08-2018-0146

Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, *6*(1), 1-55. https://doi.org/10.1080/10705519909540118

Huang, X., Lin, C. H., & Chi Kin Lee, J. (2020). Moving beyond classroom teaching: A study of multidimensional teacher self-efficacy on job satisfaction and occupational commitment. *Teachers and Teaching*, *26*(7-8), 522-542. https://doi.org/10.1080/13540602.2021.1890014

Imants, J., & Van der Wal, M. M. (2020). A model of teacher agency in professional development and school reform. *Journal of Curriculum Studies*, *52*(1), 1-14. https://doi.org/10.1080/00220272.2019.1604809

Karasar, N. (2020). *Bilimsel araştırma yöntemi.* Nobel Yayın.

Karaoğlu, İ. B. (2019). Öğretmen öz yeterlik algısı ölçeği kısa formu'nun Türkçe'ye uyarlanması: Geçerlik ve güvenirlik çalışması. *Akademik Sosyal Araştırmalar Dergisi*, *7*(99), 123-139. https://doi.org/10.29228/ASOS.36797

Keskin, T., & Aktay, S. (2021). Sınıf öğretmenlerinin eğitsel internet kullanım öz-yeterlikleri ve öğretmenlik öz-yeterliklerinin incelenmesi. *Muğla Sıtkı Koçman Üniversitesi Eğitim Fakültesi Dergisi*, *8*(2), 579-596. https://doi.org/10.21666/muefd.846281

Kılınç, A. Ç., Polatcan, M., & Çepni, O. (2023a). Exploring the association between distributed leadership and student achievement: The mediation role of teacher professional practices and teacher self-efficacy. *Journal of Curriculum Studies*, *55*(3), 352-368. https://doi.org/10.1080/00220272.2023.2216770

Kılınç, A. Ç., Polatcan, M., Erdoğan, O., Sezgin, F., & Özdemir, S. (2023b). Investigating the association between principal learning-centred leadership and teacher instructional practices: The mediating roles of teacher self-efficacy and collective teacher efficacy. *Educational Studies*, *51*(4), 568-587. https://doi.org/10.1080/03055698.2023.2279035

Kim, L. E., Dar-Nimrod, I., & MacCann, C. (2018). Teacher personality and teacher effectiveness in secondary school: Personality predicts teacher support and student self-efficacy but not academic achievement. *Journal of Educational Psychology*, *110*(3), 309-323. https://doi.org/10.1037/edu0000217

Kline, R. B. (2023). *Principles and practice of structural equation modeling*. Guilford Publications.

Korthagen, F. (2017). Inconvenient truths about teacher learning: Towards professional development 3.0. *Teachers and Teaching*, *23*(4), 387-405. https://doi.org/10.1080/13540602.2016.1211523

Kwakman, K. (2003). Factors affecting teachers' participation in professional learning activities. *Teaching and Teacher Education*, *19*(2), 149-170. https://doi.org/10.1016/S0742-051X(02)00101-4

Lai, C., Li, Z., & Gong, Y. (2016). Teacher agency and professional learning in cross-cultural teaching contexts: Accounts of Chinese teachers from international schools in Hong Kong. *Teaching and Teacher Education*, (54), 12-21. https://doi.org/10.1016/j.tate.2015.11.007

Lee, S. W., & Lee, E. A. (2020). Teacher qualification matters: The association between cumulative teacher qualification and students' educational attainment. *International Journal of Educational Development*, (77), 102218. https://doi.org/10.1016/j.ijedudev.2020.102218

Lemon, N., & Garvis, S. (2016). Pre-service teacher self-efficacy in digital technology. *Teachers and Teaching*, *22*(3), 387-408. https://doi.org/10.1080/13540602.2015.1058594

Li, L., & Ruppar, A. (2021). Conceptualizing teacher agency for inclusive education: A systematic and international review. *Teacher Educationand Special Education*, *44*(1), 42-59. https://doi.org/10.1177/0888406420926976

Lieberman, A., & Pointer Mace, D. (2010). Making practice public: teacher learning in the 21st century. *Journal of Teacher Education*, *61*(1-2), 77-88. https://doi.org/10.1177/0022487109347319

Lindell, M. K., & Whitney, D. J. (2001). Accounting for common method variance in cross-sectional research designs. *Journal of Applied Psychology*, *86*(1), 114-121. https://doi.org/10.1037/0021-9010.86.1.114

Liu, S., Hallinger, P., & Feng, D. (2016). Supporting the professional learning of teachers in China: Does principal leadership make a difference?. *Teaching and Teacher Education*, (59), 79-91. https://doi.org/10.1016/j.tate.2016.05.023

Marschall, G., & Watson, S. (2022). Teacher self-efficacy as an aspect of narrative self-schemata. *Teaching and Teacher Education*, (109), 103568. https://doi.org/10.1016/j.tate.2021.103568

Meirink, J. A., Imants, J., Meijer, P. C., & Verloop, N. (2010). Teacher learning and collaboration in innovative teams. *Cambridge Journal of Education*, *40*(2), 161-181. https://doi.org/10.1080/0305764X.2010.481256

Meirink, J. A., Meijer, P. C., Verloop, N., & Bergen, T. C. (2009). Understanding teacher learning in secondary education: The relations of teacher activities to changed beliefs about teaching and learning. *Teaching and Teacher Education*, *25*(1), 89-100. https://doi.org/10.1016/j.tate.2008.07.003

Michos, K., Cantieni, A., Schmid, R., Müller, L., & Petko, D. (2022). Examining the relationship between internship experiences, teaching enthusiasm, and teacher self-efficacy when using a mobile portfolio app. *Teaching and Teacher Education*, (109), 103570. https://doi.org/10.1016/j.tate.2021.103570

Mifsud, C. L., & Vella, L. A. (2018). Teacher agency and language mediation in two Maltese preschool bilingual classrooms. *Language, Culture and Curriculum*, *31*(3), 272-288. https://doi.org/10.1080/07908318.2018.1504400

Min, M. (2023). School culture, self-efficacy, outcome expectation, and teacher agency toward reform with curricular autonomy in South Korea: A social cognitive approach. *Asia Pacific Journal of Education*, *43*(4), 951-967. https://doi.org/10.1080/02188791.2019.1626218

Ministry of National Education. (2024). *Eğitim istatistikleri.* https://istatistik.meb.gov.tr/

Molla, T., & Nolan, A. (2020). Teacher agency and professional practice. *Teachers and Teaching*, *26*(1), 67-87. https://doi.org/10.1080/13540602.2020.1740196

Muthén, L. K., & Muthén, B. (2017). *Mplus user's guide.* Muthén & Muthén.

Özdemir, M., Kaymak, M. N., & Çetin, O. U. (2023). Unlocking teacher potential: The integrated influence of empowering leadership and authentic leadership on teacher self-efficacy and agency in Turkey. *Educational Management Administration & Leadership*, *53*(5), 1083-1104. https://doi.org/10.1177/17411432231217136

Pang, J. (2011). Exploring Korean teacher classroom expertise in sociomathematical norms. In Y. Li & G. Kaiser (Eds.), *Expertise in mathematics instruction: An international perspective* (pp. 243-262). Springer. https://doi.org/10.1007/978-1-4419-7707-6_12

Preacher, K. J., & Hayes A. F. (2008). Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models. *Behavior Research Methods,* (40), 879-891. https://doi.org/10.3758/BRM.40.3.879

Podsakoff, P. M., MacKenzie, S. B., & Podsakoff, N. P. (2012). Sources of method bias in social science research and recommendations on how to control it. *Annual Review of Psychology,* (63), 539-569. https://doi.org/10.1146/annurev-psych-120710-100452

Polatcan, M. (2020). The adaptation of professional learning activities scale to Turkish: The validity and reliability study. *International Journal of Contemporary Educational Research*, *7*(2), 281-290. https://doi.org/10.33200/ijcer.808721

Polatcan, M. (2021). An exploration of the relationship between distributed leadership, teacher agency, and professional learning in Turkey. *Professional Development in Education*, *50*(4), 789-803. https://doi.org/10.1080/19415257.2021.1905050

Polatcan, M., Arslan, P., & Balci, A. (2023). The mediating effect of teacher self-efficacy regarding the relationship between transformational school leadership and teacher agency. *Educational Studies*, *49*(5), 823-841. https://doi.org/10.1080/03055698.2021.1894549

Rani, R., Yuliasri, I., Mujiayanto, J., & Astuti, P. (2023). Enhancing teacher professional development: Insight from teacher professional learning activities. *International Journal of Social Science and Education Research Studies*, *3*(10), 2078-2085. https://doi.org/10.55677/ijssers/V03I10Y2023-15

Rockoff, J. E. (2004). The impact of individual teachers on student achievement: Evidence from panel data. *American Economic Review*, *94*(2), 247-252. https://doi.org/10.1257/0002828041302244

Sağır, M., & Tutkun, B. (2017). Okul müdürlerinin etik liderlik davranışları ile öğretmen öz yeterliği arasındaki ilişki. *Düzce Üniversitesi Sosyal Bilimler Dergisi*, *7*(2), 44-68.

Sak, İ. Ş. (2015). Okul öncesi öğretmen adaylarının sınıf yönetimi ile ilgili öz-yeterlik inançları. *Sakarya Üniversitesi Eğitim Fakültesi Dergisi*, (29), 101-120.

Seligman, M. (2018). *The optimistic child: A revolutionary approach to raising resilient children*. Nicholas Brealey Publishing.

Shavard, G. (2022). Teacher agency in collaborative lesson planning: Stabilising or transforming professional practice?. *Teachers and Teaching*, *28*(5), 555-567. https://doi.org/10.1080/13540602.2022.2062745

Supovitz, J. A., & Turner, H. M. (2000), The effects of professional development on science teaching practices and classroom culture. *Journal of Research in Science Teaching, 37*(9), 963-980. https://doi.org/10.1002/1098-2736(200011)37:9<963::AID-TEA6>3.0.CO;2-0

Şeref, İ., & Çinpolat, E. (2021). Türkçe öğretmenlerinin öğretim yöntem ve tekniklerini kullanma becerileri ile öğretmen öz yeterlikleri arasındaki ilişki. *Ana Dili Eğitimi Dergisi*, *9*(4), 1051-1065. https://doi.org/10.16916/aded.907876

Tao, J., & Gao, X. (2017). Teacher agency and identity commitment in curricular reform. *Teaching and Teacher Education*, (63), 346-355. https://doi.org/10.1016/j.tate.2017.01.010

Toom, A., Pyhaltö., & Rust, F. O. (2015). Teachers' professional agency in contradictory times. *Teachers and Teaching, 21*(6), 615-623. https://doi.org/10.1080/13540602.2015.1044334

Tschannen-Moran, M. & Hoy, A. W. (2001). Teacher efficacy: Capturing an elusive construct. *Teaching and Teacher Education*, *17*(7), 783-805. https://doi.org/10.1016/S0742-051X(01)00036-1

Tschannen-Moran, M., & McMaster, P. (2009). Sources of self-efficacy: Four professional development formats and their relationship to self-efficacy and implementation of a new teaching strategy. *The Elementary School Journal*, *110*(2), 228-245. https://doi.org/10.1086/605771

Tschannen-Moran, M., Woolfolk-Hoy, A., & Hoy, W. K. (1998). Teacher efficacy: Its meaning and measure. *Review of Educational Research, 68*(2), 202-248. https://doi.org/10.3102/00346543068002202

Wang, H., Hall, N. C., & Rahimi, S. (2015). Self-efficacy and causal attributions in teachers: Effects on burnout, job satisfaction, illness, and quitting intentions. *Teaching and Teacher Education*, (47), 120-130. https://doi.org/10.1016/j.tate.2014.12.005

Wild, A., Galosy, J., Kagle, M., Gillespie, N., & Rozelle, J. (2018). Teacher agency over curriculum and professional learning: Lock-step. *Journal of Professional Capital and Community*, *3*(4), 306-320. https://doi.org/10.1108/JPCC-12-2017-0034

Woodcock, S., Sharma, U., Subban, P., & Hitches, E. (2022). Teacher self-efficacy and inclusive education practices: Rethinking teachers' engagement with inclusive practices. *Teaching and Teacher Education*, (117), 103802. https://doi.org/10.1016/j.tate.2022.103802

World Bank. (2022). *Teachers*. https://www.worldbank.org/en/topic/teachers#1

Voogt, J., Westbroek, H., Handelzalts, A., Walraven, A., McKenney, S., Pieters, J., & De Vries, B. (2011). Teacher learning in collaborative curriculum design. *Teaching and Teacher Education*, *27*(8), 1235-1244. https://doi.org/10.1016/j.tate.2011.07.003

Zee, M., & Koomen, H. M. (2016). Teacher self-efficacy and its effects on classroom processes, student academic adjustment, and teacher well-being: A synthesis of 40 years of research. *Review of Educational Research*, *86*(4), 981-1015. https://doi.org/10.3102/0034654315626801

Education and Science

Original Article

Vol 51 (2026) No 226 461-479

The glass ceiling syndrome: research in the context of school administrators

Ayhan Kandemir [1]

Abstract

The purpose of this study is to reveal school administrators' views on the glass ceiling syndrome. The phenomenological method, which is among qualitative methods, was used in the study. The study group consisted of ten female and ten male school administrators working in public schools in the center of Bolu province during the first semester of the 2023-2024 academic year, selected through convenient sampling method. The study data were collected using a semi-structured interview form created by the researcher. The data obtained were analyzed using descriptive analysis and content analysis, and the data were presented as frequencies based on repetition. The results of the study showed that while female administrators expressed negative views regarding their development in the field of management due to taking on multiple roles, male administrators expressed both negative and positive views. It was found that both female and male administrators generally expressed neutral or positive views regarding the impact of organizational factors on the development of female administrators in management. It was observed that both female and male administrators perceived female administrators as successful in some situations and unsuccessful in others, and expressed the view that working with female administrators had both advantages and disadvantages. Finally, it was concluded that both female and male administrators have different views on the problems female administrators face due to their gender. In light of these results, recommendations were made, such as conducting similar studies at the senior management level (branch/district/provincial administrators) and implementing more positive discrimination in administratorial appointments for women.

Keywords

Glass ceiling syndrome

Female

School

Teacher

Administrators

Article Info

Received: 05.04.2024

Accepted: 10.09.2025

Published Online: 04.30.2026

DOI: 10.15390/ES.2026.2500

Introduction

It is well known that schools, which are among the important organizations of our time, are institutions that play a leading role in achieving the educational goals of states. It goes without saying that it is important for schools to have effective management in order to achieve the educational goals set by the state. In this regard, it is considered important to attach importance to the professional competencies of school administrators who perform administrative duties in schools and to respect the

[1] Bolu Atatürk Middle School, Bolu, Türkiye, ayh_81@hotmail.com

rights of administrators, such as personal and professional development. However, research suggests that administrators today face major problems, especially with regard to their professional development, and that female administrators in particular face obstacles due to their gender. It is argued that the concept of the "glass ceiling syndrome" has an important place in revealing these obstacles.

The glass ceiling syndrome was first mentioned by C. Hymowitz and T. D. Schellhardt in the Wall Street Journal in 1986 and was widely considered to be the barriers women face in hierarchical promotion (Jackson, 2014; Weyer, 2007). In 1987, Morrison et al. (1987, cited in Atacan & Genç, 2023) used the concept of the glass ceiling as a transparent barrier that prevents women from advancing beyond a certain level in organizations. In other words, the glass ceiling is defined as an artificial barrier that prevents qualified women working in the organization from advancing to leadership positions (Babic & Hansez, 2021; Jackson & O'Callaghan, 2009; Jefferson, 2019; Macarthur & Samblanet, 2010; Mert, 2019). The glass ceiling is based on the fact that women see the goals they set for themselves in their careers, but encounter various problems in achieving these goals (Yamagata et al., 1997). Based on the definitions in the literature, Atacan (2023, p. 40) defines the glass ceiling as invisible but difficult to overcome barriers to women's advancement to the executive ranks, attitudinal and organizational biases, and gender-based exclusions due to being a woman. Based on these definitions, it is assumed that the glass ceiling syndrome is caused by various reasons. Karaca (2007, p. 53) examined the reasons for the glass ceiling in three groups: individual, organizational and social factors. Women's personal preferences and perceptions, as well as their multiple roles, are among the individual factors. Organizational factors include organizational culture and policies, lack of mentors, and lack of participation in informal communication networks. Social factors include occupational segregation and gender-based stereotypes.

When examining definitions of the glass ceiling syndrome, it becomes clear that the glass ceiling is a concept closely related to ideas such as feminism, the glass wall, and the queen bee syndrome. Feminism, which is frequently encountered in daily life, is an approach that emphasizes the liberation of women and political, economic, cultural, and social equality between women and men by opposing gender discrimination (Sağlam, 2020; Taş, 2016). Feminism is "*A movement that seeks to end sexism, sexist exploitation, and oppression*" (Hooks, 2012, p. 9). Similarly, feminist theory opposes the presence of gender in determining opportunities in life and draws attention to the fact that it is women who suffer from gender discrimination (MacKinnon, 2020). In the concept of the glass ceiling, female employees are often directed to lower positions within management, and are generally placed in strategically insignificant positions rather than in upper management (Ryan, 2015, cited in Acar, 2024). The queen bee concept is defined as a situation in which female professionals who have achieved success in a male-dominated work environment deliberately distance themselves from their peers, making work life difficult for them (Baykal, 2018). Furthermore, the queen bee syndrome is seen as an inseparable feature for successful women, causing gender discrimination in the workplace (Sobczak, 2018). Based on these explanations, it can be seen that the concepts of feminism, the glass wall, and the queen bee syndrome are similar to the glass ceiling syndrome, emphasizing various negative aspects, primarily the problem of career advancement.

Women have different roles such as mother and wife, and they appear in different positions in business life. On the one hand, women work for the happiness and peace of their families and try to take responsibility for their children, and on the other hand, they can participate in career planning at their workplace. However, family responsibilities take precedence over career for many women and can be an obstacle to women's advancement (Akyüz, 2021; Anafarta et al., 2008; Atacan, 2023; Atacan & Genç, 2023; Atay et al., 2022; Belkıs, 2016; Yelkikalan, 2006). Individual barriers to women's advancement also include the fear that their family life will be damaged by taking on multiple roles. Women's lack of self-confidence that they will not be able to move up is seen as an important obstacle (Fettahlıoğlu & Çelik, 2007).

It is a fact that organizational culture and politics occupy an important place among the organizational factors that are among the causes of the glass ceiling syndrome. In male-centric organizational cultures, the idea that men have better leadership skills than women prevails (Mert, 2019). The lack of mentors is one of the organizational factors of the glass ceiling syndrome. Women usually want a woman mentor, but they may face difficulties in this regard because the number of female administrators is low. In addition, women at the executive level in organizations may view their successful peers as a threat and therefore be reluctant to build relationships (Douse, 2009). It is assumed that the lack of women in leadership positions reveals a lack of participation in informal communication networks. Women, of whom there are few in leadership positions, do not find enough space for themselves in the communication networks, and therefore their social circles are not as large as those of men. While men have strong communication with each other, it is assumed that they encounter women with a negative social image (Bilimoria & Liang, 2007, cited in Mert, 2019). The fact that women have more family responsibilities in addition to their work can also have a negative impact on women's ability to engage after hours and thus interact with administrative staff (Jackson, 2001).

It is assumed that occupational segregation and gender stereotypes occupy an important place among the social factors that are considered to be a further cause of the glass ceiling syndrome. In the shadow of the traditional social structure, the idea arises that men will be more successful in management positions than women. Women are exposed to the glass ceiling due to this understanding, and as a result, they may have difficulty advancing their careers (Atacan, 2023). As a result of this obstacle, women may have to spend more time and effort (Karaca, 2007). It is argued that women possess characteristics such as being emotional, gentle, and helpful, while men possess characteristics such as being aggressive, confident, and ambitious. Therefore, leadership roles are often perceived as more suitable for men (Duehr & Bono, 2006). Based on these findings, it can be said that various reasons cause the glass ceiling syndrome in women and that these reasons can occur in all organizations. From this point of view, it can be interpreted that women in schools, which are important organizations in society, may have problems in advancing to higher management positions due to the glass ceiling syndrome.

The literature indicates that the glass ceiling syndrome has significant consequences in addition to its causes. Taparia and Lenka (2022) stated that the glass ceiling syndrome leads to consequences such as job satisfaction, burnout, occupational stress, decreased productivity, low organizational performance, and increased employee turnover. Babic and Hansez (2021) found in their studies that the glass ceiling triggers work-family conflict, which in turn increases the likelihood of leaving the job and reduces job commitment. Similarly, Wei et al. (2025) found in their study in China that the glass ceiling syndrome affects burnout among women and has an impact on work-family conflict and burnout. It is evident that the glass ceiling syndrome causes both individual and organizational problems in this context.

The glass ceiling syndrome can be linked to leadership theories. The findings indicate that an effective leadership style in organizations is an important factor in preventing the glass ceiling syndrome. For example, transformational leaders coach their followers, creating an environment for their personal development and learning (Harms & Crede, 2010). Ethical leaders, on the other hand, prioritize the interests of their group rather than their own, respect the rights of employees in accordance with ethical principles, and treat them fairly (Zhu et al., 2004, cited in Esmer, 2013). From this, it can be concluded that the effects of the glass ceiling syndrome can be minimized in an organization with effective leadership characteristics. This can be attributed to a negative relationship has been observed between school leadership and glass ceiling barriers (Kozan et al., 2025). Considering that the glass ceiling syndrome is thought to be effective on concepts such as organizational justice (Tunç & Özmen, 2016), organizational trust (Örücü & Akgül, 2019), job satisfaction (Yıldız & Bedük, 2020), and job fulfillment (Duruk & İnce, 2024).

The glass ceiling syndrome is seen to be not just a problem for one country or region, but rather one of the common problems facing the world in general. For example, Dimovski et al. (2010) found in their study conducted in Singapore and Malaysia that female administrators encountered the glass ceiling and lacked sufficient institutional support. Joseph and Shaji (2020) identified the main glass ceiling barriers in their study in India as family responsibilities and organizational practices, gender stereotypes, organizational culture, and individual barriers. Achour (2025) concluded in his study in Tunisia that there is a positive relationship between perceived glass ceiling and burnout. Sharma and Kaur (2019) concluded in their study in India that female administrators face the glass ceiling in terms of organizational and social barriers. In the United Kingdom, gender discrimination is a problem in the workplace, and although there has been progress, the needle still points towards men in senior positions (Nationwide Employment Lawyers, 2025). Livingstone et al. (2016), in their study conducted in Canada, found that glass ceilings perpetuated by men and women's primary responsibility for household chores are the biggest obstacles to equitable promotion.

In Turkey, there are more women than male teachers in schools. For example, 409.063 of the 975.698 teachers working in public education institutions in the 2021-2022 academic year are male and 566,635 are women (Ministry of National Education [MoNE], 2023). On the other hand, the proportion of women in leadership positions, including school administrators, is of course significantly lower than that of men. According to the Directorate General for the Status of Women (2023, pp. 28-29), 10.56% (3438) of school principals, 10.24% (282) of deputy principals and 25.35% (14032) of assistant principals are women. In the provincial organization of the Ministry of National Education (MoNE), 3.7% (3) of national education directors at provincial level, 1.46% (11) of permanent national education directors at district level, 4.88% (14) of deputy national education directors at provincial level and 7.4% (199) of permanent branch directors are women. This data shows that while there are more women teachers than male teachers in Turkey, the opposite is true for administrators. It is known that the glass ceiling syndrome is an important obstacle for women to make a career in any field, especially in management. In Turkey, it can be assumed that the glass ceiling syndrome is the reason why women teachers are not sufficiently involved in higher management, especially in school administration. Considering that there are many different reasons for the glass ceiling syndrome, it is important to uncover these reasons. The aim is therefore to make suggestions to the relevant decision-makers on how to solve this problem and to contribute to the relevant literature.

Purpose of the Study

The purpose of this study is to determine the views of school administrators on the glass ceiling syndrome. To this end, female administrators were asked the following questions:

1-How would you rate the impact of taking on more than one role as a woman administrator (mother, wife, administrator, etc.) on your development in the field of administration? 2-How would you rate the impact of organizational factors (organizational culture, organizational policies, inequality of opportunity, communication issues, etc.) on your development in the field of administration as a women administrator? 3-How would you rate the success of female administrators in the field of school administration as a female administrators? 4-As a women administrator, what do you think about working with women administrators? (advantages/disadvantages) 5-As a women administrator, should there be affirmative action in favor of female administrators? What do you think about this issue? 6-What difficulties do you face in your life as a women administrator? What are your views on this issue?

The male administrators were asked the following questions: 1- How do you assess the impact of female administrators taking on more than one role (mother, wife, administrator, etc.) on their development in the field of management? 2- In your opinion, what impact do organizational factors (organizational culture, organizational policies, inequality of opportunity, communication problems, etc.) have on the development of female administrators in the field of administration? 3- How do you assess the success of female administrators in the field of school administration? 4- What do you think about working with female administrators? (advantages/disadvantages) 5- What do you think of positive discrimination in favor of women administrative staff? 6- What difficulties do female administrators face in their administratorial life because of their gender? What is your opinion on this topic?

Method

Research Model

In this study, which aims to determine the views of school administrators about the glass ceiling syndrome, the phenomenological method, one of the qualitative research methods, was used. Qualitative research is known as an approach in which verbally and visually detailed data is collected, analyzed, and presented to uncover situations, facts, events, and relationships (Gay et al., 2012). Phenomenological design is known as a research method (Yıldırım & Şimşek, 2011) that focuses on phenomena that are recognized but lack deep understanding. As the aim was to uncover school administrators' views on blanket syndrome in detail, the phenomenological design was used.

Study Group

The study group of the current research consists of 10 women and 10 male administrators (principals/vice principals) working at different levels of public education in Bolu Center in the first semester of the 2023-2024 academic year. Instead of the maximum diversity sampling method, an convenient sampling method was used in sample selection. In maximum diversity sampling, the aim is to maximize the demographic diversity of the individuals whose views and experiences regarding the problem under investigation will be examined (Yıldırım & Şimşek, 2011). However, due to the limited number of female administrators available under the current conditions, which reduces diversity, an appropriate sampling method was employed in the study. A convenient sampling method is one that facilitates reaching and including individuals or groups within the scope of the research (Ekiz, 2009). On the other hand, the fact that researchers cannot predict the representativeness of the sample (Nachimas & Nachimas, 1996, cited in Yıldız, 2017) is seen as a limitation of the appropriate sampling method. In this study, importance was placed on participants being volunteers and having worked or currently working with administrators of the opposite sex, thus paving the way for participants to give more realistic answers on the subject.

Table 1. Demographic Characteristics of the Participants

Participant	Total Length of Service in Management	Type of School	Education Level	Participant	Total Length of Service in Management	Type of School	Education Level
FA1	Over 21 years	Secondary School	Bachelor's degree	**MA1**	13-16 years	High School	Bachelor's degree
FA2	9-12 years	Primary school	Master's degree	**MA2**	13-16 yeras	Primary school	Bachelor's degree
FA3	13-16 years	High School	Bachelor's degree	**MA3**	9-12 years	Secondary School	Bachelor's degree
FA4	5-8 years	High School	Master's degree	**MA4**	17-20 years	Secondary School	Master's degree
FA5	9-12 years	Secondary School	Master's degree	**MA5**	9-12 years	Preschool	Bachelor's degree
FA6	9-12 years	Secondary School	Master's degree	**MA6**	13-16 years	High School	Master's degree
FA7	5-8 years	Secondary School	Master's degree	**MA7**	13-16 years	Preschool	Bachelor's degree
FA8	5-8 years	Preschool	Bachelor's degree	**MA8**	13-16 years	High School	Master's degree
FA9	0-4 years	Preschool	Bachelor's degree	**MA9**	13-16 years	Preschool	Master's degree
FA10	Over 21 years	Primary school	Bachelor's degree	**MA10**	Over 21 years	Secondary School	Bachelor's degree

Table 1 shows the demographic characteristics of the participants. When the table is analyzed, it is seen that 10 of the participants are women, and 10 of them are male. It was also seen that 1 of the administrators had 0-4 years of seniority, 3 of them had 5-8 years of seniority, 5 of them had 9-12 years of seniority, 7 of them had 13-16 years of seniority, 1 of them had 17-20 years of seniority, and 3 of them had 21 years or more of seniority. In addition, 5 of the administrators worked in preschool, 3 in primary school, 7 in secondary school, and 5 in high school; 11 of them had bachelor's degrees, and 9 of them had master's degrees.

Data Collection Tool

School administrators' views on the glass ceiling syndrome were elicited using a semi-structured interview form. The interview, which provides the opportunity to learn about individuals' knowledge, thoughts, attitudes and behaviors on various issues and the possible reasons for these issues in the shortest possible way (Karasar, 2018), was considered suitable for the present research. To this end, the interview questions were created by reviewing the national and international literature to determine the thoughts of school administrators about the glass ceiling syndrome in a brief form. In the validity and reliability phase, the opinions of three faculty members specializing in the field were obtained, thereby aiming to ensure internal validity. In the next phase, a pilot study was conducted with two women and two male administrative employees who were not part of the study group to determine whether the interview questions were clear and understandable for the participants.

Data Collection

The data required for the study were collected in accordance with the ethical committee approval obtained from the Human Research Ethics Committee of Bolu Abant İzzet Baysal University at its meeting dated October 12, 2023, and numbered 2023/08. The required data were collected in person by the researcher in December and January by obtaining consent from the volunteer administrators. Prior to the interview, the administrators were informed about the study and advised that their open

identities would not be included in the study. This enabled the administrators to be more open and objective about the topic. No fixed time constraint was imposed on the interviews, and in general the interviews lasted around 30 minutes.

Data Analysis

The present study was conducted using qualitative data analysis. The main purpose of qualitative data analysis is to uncover hidden information within the social structure. The set of studies in which the issues related to the study are revealed after separating and classifying the data obtained through techniques such as observation and interview and the studies are reported is called qualitative data analysis (Özdemir, 2010). In the present study, which aims to reveal the views of administrative employees about the glass ceiling syndrome, the required data were analyzed using the methods of descriptive analysis and content analysis. The purpose of content analysis is to guide the academic studies planned for the future related to the subject under investigation and to reveal the general trend situation regarding the subject (Ültay et al., 2021). Descriptive analysis, on the other hand, is a type of analysis in which the data obtained are regularly interpreted and communicated to the researchers, usually with verbatim quotations, as the opinions obtained are presented as a whole (Yıldırım & Şimşek, 2011). For this purpose, the opinions obtained in the current study have been coded and presented. The most frequently repeated statements have been grouped under specific headings and shown in tables. The tables are sorted from the most frequent opinion to the least frequent. The opinions have also been classified internally, with the aim of providing the reader with detailed results. In addition, to enable the reader to access detailed information, the themes have been supported by different statements that constitute the theme and presented according to their frequency. For example, participants' views on the impact of female administrators taking on multiple roles on their development in the field of management have been classified under the themes of no impact, positive impact, and negative impact. Under the theme of negative impact, the views of female and male participants are also presented in the table according to frequency. The researcher carefully read and interpreted the opinions obtained in order to minimize potential biases in the data analysis process. The aim was to minimize potential bias by including direct quotations from the participants' views. To increase the credibility of the study, the findings were compared, and to this end, a different researcher was asked to interpret the data obtained. The data analyzed by the researchers were compared, and the degree of consistency was determined. A high degree of consistency was observed in the interpretation of the data. Data that created inconsistencies were discussed again, and a consensus was reached on the data that created inconsistencies. Thus, it was concluded that the data obtained were consistent with each other and were valid and reliable. According to the opinions obtained, the data were presented with frequency so that the views of the administrators were presented objectively. To ensure the confidentiality of the study participants, female administrators were coded as "FA1, FA2..." and male administrators as "MA1, MA2..." coded. The frequencies showing the repetition of opinions were not added to 100, as a participant may have more than one opinion.

Findings

In this part of the study, the results, interpretations related to the results and direct quotations are given.

The first part of the study examined the opinions of women and male administrators on the impact of female administrators taking on more than one role (e.g. mother, wife, administrator) on their development in the field of management. The opinions of the administrative staff on this topic are listed in Table 2.

Table 2. The Impact of Female Administrators Taking on more than one Role on Their Development in the Field of Management

Female Administrator	f	Male Administrator	f
No effect	4	No effect	3
It has a negative impact. Because;		**It has a positive effect. Because;**	
Workload is too much	5	They understand children better	2
I experience role confusion	4	Their fighting spirit increases	2
I am short of time	2	They manage different people more easily	1
My motivation decreases/psychology deteriorates	1	**It has a negative impact. Because;**	
		They experience role confusion	3
		High workload	2
		They are perceived as overly emotional	1

Table 2 presents the views of the administrators on the development of female administrators in the field of management when they assume more than one role. When the table is analyzed, it is seen that the opinions of female administrators are grouped under the headings of "no effect" (f=4), and negative effect. It was understood that the opinion of "workload is too much" (f=5) was expressed most frequently by female administrators. The opinions of male administrators were grouped under the headings "no effect" (f=3), "it has a positive effect", and "it has a negative effect". In the opinion that it has a positive effect; "They understand children better" (f=2), and "Their fighting spirit increases" (f=2); in the opinion that it has a negative effect; "They experience role confusion" (f=3) were the most frequently expressed opinions. When the table is analyzed, it is seen that both women, and male administrators have different opinions on the subject.

The actual opinions of the participants regarding the research are as follows: *"None at all. It did not hinder me..." (FA1), "...I have a lot of difficulty in transitioning between roles... Work intensity is the biggest obstacle in terms of my development..." (FA3), "I think it is positive. They give more of themselves to work than us. They understand children better... The fact that their work is difficult increases their fighting spirit." (MA3), "...In short, the fact that female administrators concentrate on one of their roles causes problems for them to fulfill other roles fully. They either disrupt their home or their work."* (MA10).

The second sub-problem of the study was the administrators' views on the impact of organizational factors (such as organizational culture, and policies, inequality of opportunity, communication problems) on the development of women administrators in the field of management.

Table 3. The Effect of Organizational Factors on the Development of Female Administrators in the Field of Management

Female Administrator	f	Male Administrator	f
No effect	5	No effect	4
It has a positive effect. Because;		**It has a positive effect. Because;**	
I received support	4	They receive support	2
It has a negative impact. Because;		**It has a negative impact. Because;**	
I had communication problems	1	They have communication problems	3
		They experience inequality of opportunity	2
		They are exposed to negative prejudices	1

Table 3 shows the opinions of women, and male administrators on the effect of organizational factors on the development of women administrators in the field of management. When the table is examined, it is seen that the opinions of women (f=5), and male (f=4) administrators are grouped under the views of no effect, positive, and negative effect. It was understood that female administrators expressed the views of "I received support" (f=4) in their views that it had a positive effect, and "I had communication problems" (f=1) in their views that it had a negative effect. Similarly, male administrators frequently expressed the view that "they receive support" (f=2) for the view that it has a positive effect, and "they have communication problems" (f=3) for the view that it has a negative effect.

When the table is analyzed, it is seen that male administrators also have different opinions on the subject.

The actual opinions of the participants regarding the research were as follows: *"I did not experience any obstacles or problems. This is related to the situation of the schools. There is no problem with these issues in our school." (FA7), "...I get positive feedback, and support... I do not experience inequality of opportunity. The distribution of tasks is fair." (FA6), "There is no negativity related to organizational factors. However, communication problems can have a negative impact on their development in the field of management. They may have problems in communicating with some people in the school..." (MA1), "I can say that prejudices affect the development of female administrators in this field, and that female administrators find fewer opportunities than male administrators due to prejudices." (MA8).*

The third sub-problem of the study was the administrators' views on the success of female administrators in the field of school administration.

Table 4. Achievements of Female Administrators in the Field of Management

Female Administrator	**f**	**Male Administrator**	**f**
No effect of gender	2	No effect of gender	1
Successful. Because;		**Successful. Because;**	
They can do it when they want to	4	They are organized, meticulous, and detailed	3
They are organized, meticulous, and detailed	4	They can think differently from men	3
They are emotional, and motherly	3	They communicate well with women students	1
They give importance to cooperation, and solidarity	2	They communicate well with women teachers	1
They are ambitious	1	They are more motherly	1
They understand women students better	1	**Unsuccessful. Because;**	
Unsuccessful. Because;		They are too emotional	3
Their roles are too much	1	It is difficult for them to do work outside of working hours	2
Their workload is too much	1	They are not flexible	1
They have a high sense of competition	1	They cannot look at things from a broad perspective	1
		They have problems with their fellow colleagues	1

Table 4 presents the views of women, and male administrators on the success of female administrators in the field of management. It was seen that the opinions of both women (f=2), and male administrators (f=1) were grouped under the headings "Gender has no effect", successful, and unsuccessful. In addition to this, it was understood that female administrators intensively expressed the views of "They can do it when they want" (f=4), and "They are organized, meticulous, and detailed" (f=4) under the successful title, and "Their roles are too much" (f=1), "Their workload is too much" (f=1), and "They have a high sense of competition" (f=1) under the negative title. Male administrators, on the other and, were mostly united under the successful heading; "They are organized, meticulous, and detailed" (f=3) , and "They can think differently from men" (f=3), and under the unsuccessful heading; "They are too emotional" (f=3). When the table is analyzed, it is seen that both women, and male administrators have different opinions on the subject.

The actual opinions of the participants regarding the research were as follows: *"...There are successful, and unsuccessful. It doesn't matter whether they are male or women..." (FA5), "...We do our job more lovingly , and meticulously because we come here willingly...." (FA4), "Since women are emotional, they approach students more emotionally. For this reason, I think they are better with students..." FA7), "In general, female administrators are more troubled... There may be different problems... There may be a sense of competition." (FA9), "They can think differently from us. They can put forward different options better than us. They can communicate better with women students, and women teachers." (MA2), "They may be less effective. Because they involve their emotions too much. They have sudden, and emotional approaches in big events..." (MA1).*

The opinions of female and male administrators on working with female administrators (advantages/disadvantages) constitute the fourth sub-problem of the study.

Table 5. Advantages, and Disadvantages of Working with Female Administrators

Kadın Yönetici	f	Erkek Yönetici	f
Advantage;		**Advantage;**	
Our feelings, and thoughts are similar	8	Easy communication with women teachers	7
We can communicate easily	4	Easy communication with women students	7
Solidarity is achieved	3	Effective communication with women parents	5
Disadvantage;		They are more systematic	3
None	3	They are more sensitive about cleanliness, and hygiene	1
They may be inadequate in physical work	3	They can see things from different perspectives	1
They may want less work, and responsibility	3	**Disadvantage;**	
They are too emotional	1	They may be inadequate in physical work	5
High rate of taking leave, and medical reports	1	It is difficult to work outside working hours	3
They are dominants	1	They are not cold-blooded	2
They are detail-oriented	1	High rate of taking leave, and medical reports	2
Jealousy/Envy	1	Having problems with fellow colleagues	2
It is difficult to work outside working hours	1	They are too emotional	2
		They are not flexible	1

Table 5 presents the views of female and male administrators regarding working with female administrators. When the table is analyzed, it is seen that the most frequently expressed opinions of female administrators about the advantages of working with female administrators are "We have similar feelings, and thoughts" (f=8), and "We can communicate easily" (f=4). Regarding the disadvantages, it was seen that they concentrated on the views of "None" (f=3), "They may be inadequate in physical work" (f=3), "They may want less work, and responsibility" (f=3). As for the advantages of working with female administrators, male administrators concentrated on "Easy communication with women teachers" (f=7), and "Easy communication with women students" (f=7). Regarding the disadvantages, it was understood that they mostly expressed the view that "They may be inadequate in physical work" (f=5). When the table is analyzed, it is seen that both women, and male administrators have different views on the subject.

The actual opinions of the participants on the research were as follows: *"...It is easier to get along because our feelings, and thoughts are similar." (FA5), "...They may be inadequate in repair, and renovation works. It can be difficult for them to stay late at school or work outside of working hours..." (FA8), "... There can be a lot of jealousy, and envy among women..." (FA10), "They communicate more easily with women teachers, and women students, and more effectively with women parents." (MA4), "There can be problems with physical work. It is difficult to work outside working hours. They panic more in the face of events. They are not cold-blooded." (MA3), "They are too emotional... They can be harsh in their attitudes, and behaviors, especially towards their fellow students." (MA8)*

The views of women, and male administrators on positive discrimination in favor of female administrators constituted the fifth sub-dimension of the study.

Table 6. Opinions on Positive Discrimination in Favor of Female Administrators

Female Administrator	f	Male Administrator	f
Yes. It should be done;		**Yes. It should be done;**	
In administrative appointments	7	In administrative appointments	7
In workload, and on shifts	2	In workload, and on shifts	1
No. It should not be done;		**No. It should not be done;**	
Because there should be equality between women, and men	4	Because there should be equality between women, and men	4
Because administration is a volunteer job	3	Because there must be merit	2
		Because administration is a volunteer job	1

Table 6 shows the opinions of women, and male administrators on positive discrimination in favor of female administrators. When the table is analyzed, it is seen that both women (f=7), and male (f=7) administrators expressed the opinion "In administrator appointments" most frequently under the heading "It must be done". Similarly, both women (f=4), and male (f=4) administrators frequently expressed the opinion "Because there should be equality between women, and men" under the heading "It must not be done". When the table is analyzed, it is seen that both women, and male administrators have different opinions on the subject.

The actual opinions of the participants regarding the research are as follows*: "...It should be a must in appointments. There should be a woman administrator in every school. Because women teachers, and women students can explain their situations better to female administrators..." (FA2), "There should be more female administrators. In some schools, it may be compulsory to have only women administrators. It should not be during working hours. Women, and men should be equal. Because being an administrator is a voluntary job." (FA8), "Measures should be taken to have female administrators in schools... Working hours, and workload should be the same. Because this situation can be abused." (MA7), "No, it should not be. You should give the job to those who know the job. It is important that people with merit come. Gender is not important." (MA9)*

The sixth sub-dimension of the study is the views of women, and male administrators on the problems faced by female administrators in their administratorial lives due to their gender.

Table 7. Problems Faced by Female Administrators due to Their Gender

Female Administrator	**f**	**Male Administrator**	**f**
Yes, we do;		**Yes, they do;**	
Communication problems with parents	4	Rude behavior, and bullying from parents	3
Role confusion	3	Communication problems	3
Work overload	2	Work overload	2
Lack of time for ourselves, and our families	2	Disciplinary problems for students	2
Students use our goodwill	1	Role confusion	2
Difficulty with physical work	1	Social norms	2
No, we don't. Because;		negative prejudices	1
Our attitude, and stance are clear	4	**No, they do not. Because;**	
We are experienced	3	They are privileged because they are women	2
Our working environment is harmonious	1	Our parents have a high level of culture	1
		Their attitudes are clear	1

Table 7 shows the opinions of women, and male administrators on the problems that female administrators face in the field of administration due to their gender. It was observed that female administrators tors expressed the views of "Communication problems with parents" (f=4) most frequently under the yes heading, and "Our attitude, and stance are clear" (f=4) most frequently under the no heading. Male administrators, on the other h, and, mostly expressed "Rude behavior, and bullying from parents" (f=3) , and "Communication problems" (f=3) under the yes heading, and mostly expressed the opinion that "They are privileged because they are women" (f=2) under the no heading. When the table is analyzed, it is seen that both women, and male administrators have different views on the subject.

The actual opinions of the participants regarding the research are as follows*: "We encounter problems. There are too many roles, and duties on women. There can be problems with physical work. We may have problems with parents who are not underst, anding." (FA5), "I don't have problems because my working environment is harmonious. We do not have problems because we are experienced, and st, and upright." (FA9), "They do. It can be seen in disciplinary problems for students..." (MA3), "They may experience rude behavior, and bullying from parents. Social norms, and prejudices are also effective in this. Of course, the environment where the school is located is also important." (MA2), "...No. Because they are few in number. Therefore, they are privileged because they are women." (MA1)*

Conclusion, Discussion, and Recommendations

This part of the study presents the results, the discussion and the proposals related to the results.

As part of the first sub-problem of the study, the opinions of female and male administrators on the development of female administrators in the field of management due to taking on more than one role were included. Although there were opinions among female and male administrators that this situation has no impact, it was found that there were also opinions expressing that female administrators face negative aspects such as high workload and role confusion. This can be attributed to women having to exert greater effort due to taking on different roles, the simultaneous emergence of work and family demands, or women experiencing conflict due to role incompatibility (Akyüz, 2021). Additionally, organizational culture and policies may also be considered influential in this situation. Indeed, male-dominated organizational culture contributes to the glass ceiling being felt more strongly (Karcıoğlu & Leblebici, 2014). In such work environments, it is difficult for women to advance to leadership positions. To eliminate this, cultural change and policies and practices that support gender equality must be developed (Northouse, 2021, cited in İş & Solmaz, 2025). In addition, male administrators also expressed the view that the fact that administrators take on more tasks leads to positive outcomes, such as a better understanding of children and an increase in their fighting spirit. When examining the relevant literature, it was found that there are results that support the results of the current study. For example, Mert (2019) examined the barriers that prevent female teachers from becoming administrators in the context of glass ceiling syndrome in her doctoral thesis and concluded that individual barriers include multiple roles such as mother or wife. Similarly, Atacan (2023) examined the impact of glass ceiling syndrome on the career development of female employees and found that reasons such as women taking on multiple roles and the child factor were detrimental to the promotion of female administrators. Kirişçi and Can (2020) examined the views of educational and school administrators on the glass ceiling syndrome and concluded that the problems of women such as mothers, wives, etc. strongly affect the performance of administrators. In the same study, it was found that some participants thought that this situation was not effective and the result was similar to the research results. When the literature is examined, it is seen that there are various studies (Aksu et al., 2013; Bulut & Çelikten, 2021; Kalafatoğlu & Torun, 2022; Liff & Ward, 2001; Özkan et al., 2023; Sari, 2012; Yücedağ & Günbayı, 2016) that are similar to the research results.

The second sub-problem of the study examined the views of female and male administrators on the impact of organizational factors on the development of administrators in the area of management. Both female and male administrators overwhelmingly believed that organizational factors were not effective in the development of administratorial administrators. They also overwhelmingly believed that organizational factors have a positive impact on women's development and that women are supported. In addition, it was observed that some male administrators in particular expressed that women in administration face organizational barriers such as communication problems and inequality of opportunity. In general, it can be concluded that there is a positive organizational environment for female administrators in schools and that the development of women in their administrative careers is supported within the organization. Aksu et al. (2013) concluded in their studies that both female and male school administrators are not significantly hindered in terms of organizational culture. Similarly, İş and Solmaz (2025) concluded in their article on female administrators that female administrators encounter obstacles such as inequality of opportunity and organizational biases in their path to becoming administrators, which is consistent with the research results. However, there are also studies that differ from the results of the research. For example, Atacan's (2023) study found that women in administration face obstacles such as bullying, male dominance in management, gender discrimination, hiring male administrators, and women being seen as unsuitable for leadership in the context of organizational culture and politics, which differs from the research results. One reason for this difference may be gender norms. Gender norms define the limits of people's roles and responsibilities in their lives, and social perceptions are shaped by the values and norms of

societies. Values and norms shaped by biological gender within society form the basis of individuals' attitudes within that society. Furthermore, these values and norms also shape expectations of individuals (Demir, 2020). From this, it can be concluded that the effects of gender norms continue to be seen in different studies, and that high-level positions such as management are not considered suitable for women.

The third sub-problem of the study was the views of female and male administrators on the success of female administrators in the field of school administration. There are opinions among male, and female administrators that gender has no effect on success. In addition, both male and female administrators stated that they consider female administrators successful because they are organized and meticulous. In addition, female administrators often expressed the view that administrators are successful because administrators are emotional and motherly, and male administrators often expressed the view that women are successful because women think differently than men. On the other hand, female administrators consider female administrators unsuccessful because they play a big role and have a heavy workload, and male administrators consider female administrators unsuccessful because it is difficult for them to work outside of working hours. Both female and male administrators' perception that administrators are successful in the field of management is considered important for the future career development of administrators. Taş et al. (2023) in their study on preschools and Kirişçi and Can (2020) in their study on school and education administrators supported the results of the research by concluding that some administrators think that gender is not important in management. Kirişçi and Can (2020) also stated in their study that some female administrators were successful in terms of sacrificial and diligent work in the field of management, successful in terms of communication and unsuccessful in terms of emotional decision making and supported the results of the current research. It is considered significant that male administrators perceive female administrators as successful in some situations and unsuccessful in others. This may be attributed to the experiences of male administrators. It can be argued that a male administrator who has worked with a female administrator who loves her job and balances her work and family life, either in the past or present, will generally view the female administrator he works with as successful. This is because the male administrator may have a positive opinion of the female administrator due to reasons such as her viewing events from a different perspective, being nurturing, and being able to connect more easily with female students. On the other hand, it is normal for a male administrator who has worked with or is working with a woman who wants to be a administrator but cannot balance her work and family life, is not supported by her family, tries to take on family responsibilities alone, etc., to have more negative opinions. In this case, it can be thought that the male administrator, who has to take on more management responsibilities at school, expresses negative opinions about female administrators.

The fourth sub-problem of the study was the opinions (advantages/disadvantages) of female and male administrators about working with female administrators. While female administrators cited similar feelings, thoughts and ease of communication as advantages of working with female administrators, they cited inadequacies in physical work, less responsibility, etc. as disadvantages. Male administrators, on the other hand, mentioned ease of communication between female administrators, teachers, parents and students as advantages and inadequacies in physical work and difficulty in working outside of working hours as disadvantages. Apart from these, both female and male administrators had different opinions (see Table 5). Among the opinions, the high number of opinions on the advantages of working with female administrative staff is considered important. This situation shows that both genders have a positive attitude towards female administrators. In their study on education and school employees, Kirişçi and Can (2020) concluded that administrators want to work with female administrators because women communicate easily, work harmoniously and think in detail. In the same study, some administrators stated that they did not want to work with women because of their emotional decision making, selfishness and moodiness, which is similar to the results of the current study. The fact that women have an advantage in communicating with students and parents due to their emotional depth (Günsel et al., 2015) could lead to women having an advantage in management communication. It can be interpreted that this situation may contribute to the success of female administrators in general.

The fifth sub-problem of the study was the views on affirmative action in favor of female administrators. Both male and female administrators expressed their opinions on affirmative action in favor of women, especially in appointments. In addition, there were also opinions about positive discrimination in favor of female administrators in terms of workload, albeit to a lesser extent. On the other hand, it was concluded that there should be no discrimination between male and female administrators for reasons such as gender equality, merit and the fact that administration is a voluntary activity. Although the number of female teachers in Turkey is high (MoNE, 2023), the number of female administrators is quite low (Directorate General for the Status of Women, 2023). Although the Ministry of Education's Regulation on the Appointment of Administrative Employees (Millî Eğitim Bakanlığına Bağlı Eğitim Kurumlarına Yönetici Seçme ve Görevlendirme Yönetmeliği, 2021) provides for positive discrimination in favor of administrative employees in some schools, this is obviously not enough. It is known that the percentage of female administrators in education is also low in Western countries, although not as low as in Turkey (İnandı et al., 2009). It is well known that social norms and prejudices play an inhibiting role in the promotion of women to leadership positions in societies. It is assumed that gender stereotypes based on these prejudices have an important function (Can, 2008; Usluer, 2000). It can be assumed that educational institutions are among the institutions in which these gender stereotypes are particularly pronounced. From this point of view, it can be assumed that these stereotypes are also effective in Turkey and prevent women from rising to the management level of schools, so that the number of female administrators in schools is low.

The sixth sub-problem of the study is the views of administrators on the difficulties women experience in their leadership lives because of their gender. Female administrators indicated that they experience problems such as communication issues with parents, role confusion, and work overload. Male administrators, on the other hand, indicated that female administrators face problems such as rude behavior, bullying by parents, communication problems, and work overload. In addition, it was found that female and male administrators also felt that female administrators did not experience any problems. While female administrators emphasized the fact that their attitudes and stances were clear and that they had experience, male administrators focused on the view that women were privileged. Özkan et al. (2023) concluded in their study on the problems of female administrators in school administration that although some female administrators stated that they had no problems, some of them had problems such as housework, the inability to allocate time for children and spouses, and the inability to communicate properly. Yücedağ and Günbayı (2016), in their study on the problems of female teachers who left their position as educational administrators in MoNE, concluded that female administrators struggled with family, social, personal, etc. Problems. It goes without saying that these two studies are comparable to the present study in terms of results. Özertürk and Gül (2021) investigated the barriers for women working in education to become administrators in the context of the glass ceiling syndrome. In the study, female administrators supported the results of the research by concluding that female administrators, unlike male administrators, had problems with parents trying to intimidate them. Furthermore, it is significant that some female administrators stated that they had no problems because their attitude and position was clear. It can be concluded that some female administrators want to continue their career despite the glass ceiling syndrome and believe that the profession of an administrator is not only suitable for men.

Studies conducted in different countries (Achour, 2025; Derin, 2020; Dimovski et al., 2010; Joseph & Shaji, 2020; Livingstone et al., 2016; Nationwide Employment Lawyers, 2025; Sharma & Kaur, 2019), it is understood that the glass ceiling syndrome is a situation generally observed in all countries. It can be concluded that the consequences of the glass ceiling syndrome, which generally arises due to individual, social, and organizational reasons, are similar worldwide and in Turkey.

When the research is evaluated overall;

While female administrators expressed negative views regarding the development of female administrators in the field of management due to their taking on multiple roles, male administrators expressed both negative and positive views.

Female and male administrators generally held neutral or positive views regarding the impact of organizational factors on the development of female administrators in the field of management.

Female and male administrators view female administrators as successful in some situations and unsuccessful in others, and believe that working with female administrators has both advantages and disadvantages.

It was found that female and male administrators have different views on the problems female administrators face due to their gender.

In light of these results, it may be recommended that similar studies be conducted in different provinces and that the research be extended to include branch administrators, district and provincial directors, in addition to school administrators. For practitioners, it may be recommended that administrators be more flexible regarding women's working hours and that seminars be held for parents, teachers, and administrators to eliminate negative preconceptions about female administrators. Finally, it may be recommended that female administrators be given positive privileges through the implementation of a quota system for women in administratorial appointments to address communication and opportunity inequality issues, and that in-service training courses be provided to improve their communication skills.

Finally, the fact that the working group was limited to the central district of Bolu and excluded employees working in different organizations, particularly private school teachers, from the scope of the research can be considered a limitation of the study

References

Acar, E. Ü. (2024). *Spor teşkilatlarında çalışan kadınlarında cam tavan sendromunun tükenmişlik ve iş performansı üzerine etkisi* (Thesis No. 915710) [Doctoral dissertation, Erciyes University]. Council of Higher Education National Thesis Center.

Achour, Z. (2025). Impact of perceived glass ceiling on burnout among female public administrators in Tunisia: the moderating role of work-life balance. *International Journal of Environment, Workplace and Employment, 9*(1), 1-15. https://doi.org/10.1504/IJEWE.2025.145494

Aksu, A., Çek, F., & Şenol, B. (2013). Kadınların müdür olmalarının önündeki cam tavan ve cam tavanı aşma stratejilerine ilişkin ilköğretim okulu müdürlerinin görüşleri. *Kocaeli Üniversitesi Sosyal Bilimler Dergisi,* (25), 133-160.

Akyüz, M. (2021). Cam tavan sendromunun bileşenlerinden çoklu rol üstlenme iş performansını etkiler mi?. *Örgütsel Davranış Çalışmaları Dergisi, 1*(1), 14-23.

Anafarta, N., Sarvan, F., & Yapıcı, N. (2008). Konaklama işletmelerinde kadın yöneticilerin cam tavan algısı: Antalya ilinde bir araştırma. *Akdeniz İİBF Dergisi, 8*(15), 111-137.

Atacan, T. (2023). *Kadın çalışanların kariyer gelişimlerinde cam tavan sendromunun etkisi: Doğu Karadeniz'deki devlet üniversitelerinin kadın akademisyenleri üzerinde bir araştırma* (Thesis No. 793921) [Doctoral dissertation, Giresun University]. Council of Higher Education National Thesis Center.

Atacan, T., & Genç, K. Y. (2023). Kadın çalışanların kariyer gelişimlerinde cam tavan sendromunun etkisi: Türkiye'de "cam tavan sendromu" konusunda yapılan lisansüstü tezlerin bibliyometrik analiz yöntemiyle incelenmesi. *Asya Studies, 7*(24), 275-298. https://doi.org/10.31455/asya.1215280

Atay, Ö., Gürgen, N. E., & İmamoğlu, E. (2022). Türkiye'de kadın akademisyenlerde algılanan cam tavan sendromu araştırması. *Ankara Üniversitesi Sosyal Bilimler Dergisi, 13*(2), 170-182. https://doi.org/10.33537/sobild.2022.13.2.15

Babic, A., & Hansez, I. (2021). The glass ceiling for women managers: Antecedents and consequences for work-family interface and well-being at work. *Frontiers in Psychology*, (12), 1-17. https://doi.org/10.3389/fpsyg.2021.618250

Baykal, E. (2018). Sosyal kimlik teorisi perspektifiyle kraliçe arı sendromu. *Karadeniz Teknik Üniversitesi Sosyal Bilimler Enstitüsü Sosyal Bilimler Dergisi, 8*(16), 159-175.

Belkıs, Ö. (2016). Anneliğin akademik kariyer gelişimine etkileri üzerine nitel bir araştırma. *Eğitim ve Öğretim Araştırmaları Dergisi, 5*(4), 250-263.

Bulut, S., & Çelikten, M. (2021). Türk eğitim sisteminde kadın okul yöneticiliği. *Journal of Social, Humanities and Administrative Sciences, 7*(40), 886-894. doi.org/10.31589/JOSHAS.621

Can, N. (2008). Okul yöneticilerinin yönetimde cinsiyet faktörüne ilişkin görüşleri. *Eğitim ve Bilim, 33*(147), 35-41. https://share.google/IIdK98jQulHEsF7T2

Demir, A. (2020). Toplumsal cinsiyet algısı ve kariyer kararsızlığı: Sinop örneği. *Avrasya Uluslararası Araştırmalar Dergisi, 8*(22), 1-17. https://doi.org/10.33692/avrasyad.689224

Derin, N. (2020). Dünyadan ve Türkiye'den örneklerle cam tavan sendromu. *Birey ve Toplum Sosyal Bilimler Dergisi, 10*(2), 137-154. https://doi.org/10.20493/birtop.814149

Dimovski, V., Skerlavaj, M., & Man, M. M. K. (2010). Comparative analysis of mid-level women managers perception of the existence of glass ceiling in Singaporean and Malaysian organizations. *International Business & Economics Research Journal, 9*(8). https://doi.org/10.19030/iber.v9i8.613

Directorate General for the Status of Women. (2023). *Türkiye'de kadın.* https://www.aile.gov.tr/media/142379/tu-rkiye-de-kadin-temmuz-23.pdf.

Douse, D. M. (2009). *Beyond the glass ceiling-a phenomenological study of senior executive women and their perceived responsibility to mentor other women* (Publication No. 3635048) [Doctoral dissertation, Capella University]. ProQuest Dissertations and Theses Global.

Duehr, E. E., & Bono, J. E. (2006). Men, women, and managers: Are stereotypes finally changing?. *Personnel Psychology, 59*(4), 815-846. https://doi.org/10.1111/j.1744-6570.2006.00055.x

Duruk, E., & İnce, F. F. (2024). Turizmde kadın çalışanların cam tavan sendromuna maruz kalma düzeyleri ile iş doyumları arasındaki ilişki: Kamu ve özel sektör işletmelerinde bir uygulama. *Türk Turizm Araştırmaları Dergisi, 8*(2), 96-114. https://doi.org/10.26677/TR1010.2024.1407

Ekiz, D. (2009). *Bilimsel araştırma yöntemleri.* Anı Yayıncılık.

Esmer, Y. (2013). Banka işletmelerinde etik liderlik ve örgütsel bağlılık ilişkisi: Ziraat Bankası örneği. *Aksaray Üniversitesi İİBF Dergisi, 5*(2), 41-61.

Fettahlıoğlu, Ö. O., & Çelik, A. (2007). Kariyer yönetiminde güncel bir kavram olarak cam tavan sendromu ve kadın yöneticiler boyutu. In M. Ş. Şimşek, A. Çelik, & A. Akatay (Eds.), *Kariyer yönetimi insan kaynakları yönetimi uygulamaları* (pp. 243-261). Gazi.

Gay, L. R., Mills, G. E., & Airasian, P. W. (2012). *Educational research: Competencies for analysis and applications.* Pearson.

Günsel, A., Köroğlu, S., & Demirci, L. (2015). Çalışma hayatında kadınların karşılaştıkları sorunlar ve cam tavan algıları: Kadın öğretmenler üzerinde bir araştırma. *KADEM Kadın Araştırmaları Dergisi, 1*(1), 74-113. https://doi.org/10.21798/kadem.2022.66

Harms, P. D., & Crede, M. (2010). Emotional Intelligence and transformational and transactional leadership: A meta-analysis. *Journal of Leadership & Organization Studies, 17*(1), 5-17. https://doi.org/10.1177/1548051809350894

Hooks, B. (2012). *Feminizm herkes içindir: Tutkulu politika* (E. Aydın, B. Kurt, Ş. Özgün, & A. Yıldırım, Trans.). BGST.

İnandı, Y., Özkan, S., Peker, S., & Atik, Ü. (2009). Kadın öğretmenlerin kariyer geliştirme engelleri. *Mersin Üniversitesi Eğitim Fakültesi Dergisi, 5*(1), 77-96.

İş, E., & Solmaz, K. (2025). Okulöncesi eğitim kurumlarında çalışan kadın yöneticilerin cam tavan sendromunun etkilerine ilişkin görüşleri. *Türkiye Sosyal Araştırmalar Dergisi, 29*(1), 91-109. https://doi.org/10.20296/tsadergisi.1359969

Jackson, J. C. (2001). Women middle managers' perception of the glass celling. *Women in Management Review, 16*(1), 30-41. https://doi.org/10.1108/09649420110380265

Jackson, J. F. L., & O'Callaghan, E. M. (2009). What do we know about glass ceiling effects? A taxonomy and critical review to inform higher education research. *Research Higher in Education,* (50), 460-482. https://doi.org/10.1007/s11162-009-9128-9

Jackson, Z. D. (2014). *A phenomenological study on organizational barriers and leadership inequalities for non-white women* (Publication No. 1721) [Doctoral dissertation, University of Phoenix]. UNLV Theses, Dissertations, Professional Papers, and Capstones. http://doi.org/10.34917/4332702

Jefferson, M. (2019). *Crystal and diamond and glass: What are women and ceilings made of?* [Doctoral dissertation, The Chicago School].

Joseph, K. J., & Shaji, P. J. (2020). Career barriers for women executives in India and the glass ceiling syndrome. *Journal of the Social Sciences, 23*(2), 552-563.

Kalafatoğlu, Y., & Torun, A. A. (2022). Kadın yöneticilerin karşılaştıkları fırsatlar ve engeller: Nitel bir çalışma. *Hacettepe Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi, 40*(3), 633-658. http://doi.org/10.17065/huniibf.1002991

Karaca, A. (2007). *Kadın yöneticilerde kariyer engelleri cam tavan sendromu üzerine uygulamalı bir araştırma* (Thesis No. 217549) [Master's thesis, Selçuk University]. Council of Higher Education National Thesis Center.

Karasar, N. (2018). *Bilimsel araştırma yöntemi* (33th ed.). Nobel.

Karcıoğlu, F., & Leblebici, Y. (2014). Kadın yöneticilerde kariyer engelleri: "Cam tavan sendromu" üzerine bir uygulama. *Atatürk Üniversitesi İktisadi ve İdari Bilimler Dergisi, 28*(4), 1-20.

Kirişçi, G., & Can, N. (2020). Eğitim ve okul yöneticilerinin cam tavan sendromuna ilişkin görüşleri. *Ahi Evran Üniversitesi Sosyal Bilimler Enstitüsü Dergisi, 6*(2), 618-636. https://doi.org/10.31592/aeusbed.736762

Kozan, A., Çoban, Ö., & Yalçın, M. T. (2025). Okul liderliği ve öğretmenlerin cam tavan engelleri arasındaki ilişkide iş birlikçi iklim. *Ege Eğitim Dergisi, 26*(1), 1-15. https://doi.org/10.12984/egeefd.1463433

Liff, S., & Ward, K. (2001). Distorted views through the glass ceiling: The construction of women's understandings of promotion and senior management positions. *Gender, Work and Organization, 8*(1), 19-36. https://doi.org/10.1111/1468-0432.00120

Livingstone, D. W., Pollock, K., & Raykov, M. (2016). Family binds and glass ceilings: Women managers' promotion limits in a 'knowledge economy'. *SAGE, 42*(1) 145-166. https://doi.org/10.1177/0896920514532663

Macarthur, K., & Samblanet, S. (2010). Gender and the glass ceiling at work. *Sociology Compass, 4*(9), 705-717. https://doi.org/10.1111/j.1751-9020.2010.00304.x

MacKinnon, C. A. (2020). *Feminist bir devlet kuramına doğru* (T. Yöney & S. Yücesoy, Trans.). Metis.

Mert, P. (2019). *Kadın öğretmenlerin yönetici olmaları önündeki engellerin cam tavan sendromu bağlamında incelenmesi* (Thesis No. 541378) [Doctoral dissertation, İstanbul Sabahattin Zaim University]. Council of Higher Education National Thesis Center.

Millî Eğitim Bakanlığına Bağlı Eğitim Kurumlarına Yönetici Seçme ve Görevlendirme Yönetmeliği (5 February, 2021). Resmî Gazete (Sayı: 31386). https://www.resmigazete.gov.tr/eskiler/2021/02/20210205-1.htm

Ministry of National Education. (2023). *Millî Eğitim istatistikleri örgün eğitim 2021/2022.* https://sgb.meb.gov.tr/www/milli-egitim-istatistikleri-orgun-egitim-20212022/icerik/460

Nationwide Employment Lawyers. (2025, 1 May). *Gender discrimination at senior levels: Are the glass ceilings thicker in 2025?.* https://natemplaw.co.uk/gender-discrimination-at-senior-levels-are-the-glass-ceilings-thicker-in-2025/

Örücü, E., & Akgül, G. (2019). Örgütlerde cam tavan sendromunun örgütsel güven üzerine etkisi: Lojistik hizmet sağlayıcılarına yönelik bir araştırma. *Ömer Halisdemir Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi, 12*(2), 183-201. https://doi.org/10.25287/ohuiibf.421932

Özdemir, M. (2010). Nitel veri analizi: Sosyal bilimlerde yöntembilim sorunsalı üzerine bir çalışma. *Eskişehir Osmangazi Üniversitesi Sosyal Bilimler Dergisi, 11*(1), 323-343.

Özertürk, M., & Gül, İ. (2021). Eğitim çalışanı kadınların yönetici olmalarının önündeki cam tavan sendromuna yönelik görüşleri. *Karadeniz Uluslararası Bilimsel Dergisi,* (51), 24-44. https://doi.org/10.17498/kdeniz.962743

Özkan, H., Demir Özkan, H., Altınok, T., & Pısıl, Y. O. (2023). Kadın yöneticilerin gözünden okul yönetiminde karşılaşılan sorunlar. *Journal of Social, Humanities and Administrative Sciences, 9*(69), 3661-3666. http://doi.org/10.29228/JOSHAS.72537

Sağlam, M. A. (2020). Feminist Kuram'da kadının politik ekonomisi ve dönüşümü. *Karadeniz Uluslararası Bilimsel Dergi,* (46), 394-412. https://doi.org/10.17498/kdeniz.717116

Sari, M. (2012). Exploring gender roles' effects of Turkish women teachers on their teaching practices. *International Journal of Educational Development, 32*(6), 814-825. https://doi.org/10.1016/j.ijedudev.2011. 08.002

Sharma, S., & Kaur, R. (2019). Glass ceiling for women and work engagement: The moderating effect of marital status. *SAGE, 8*(2) 132-146. https://doi.org/10.1177/2319714519845770

Sobczak, A. (2018). The queen bee syndrome. The paradox of women discrimination on the labour market. *Journal of Gender and Power, 9*(1), 51-61. https://doi.org/10.14746/jgp.2018.9.005

Taş, D., Taşdelen, M., Güneş, Ö., & Aybar, Y. (2023). Okul öncesi eğitim kurumlarında kadın yönetici olmak: Güçlü ve zayıf yönler, avantajlar ve dezavantajlar. *Uluslararası Eğitime Özgün Bakış Dergisi, 1*(1), 193-216, https://doi.org/10.5281/zenodo.7694980

Taş, G. (2016). Feminizm üzerine genel bir değerlendirme: Kavramsal analizi, tarihsel süreçleri ve dönüşümleri. *Akademik Hassasiyetler, 3*(5), 163-175.

Taparia, M., & Lenka, U. (2022) An integrated conceptual framework of the glass ceiling effect. *Journal of Organizational Effectiveness: People and Performance, 9*(3), 372-400. https://doi.org/10.1108/JOEPP-06-2020-0098

Tunç, S., & Özmen, A. (2016). Kadın çalışanların cam tavan engelleri ile algılanan örgütsel adalet arasındaki ilişki: Eskişehir bankacılık sektöründe uygulama. *ÇKÜ Sosyal Bilimler Enstitüsü Dergisi, 7*(2), 97-119.

Usluer, L. (2000). *Kadın öğretmenlerin yönetici konumlara yükseltilmeme nedenleri konusundaki öğretmen ve yönetici görüşlerinin değerlendirilmesi (Ankara ili örneği)* (Thesis No. 94732) [Master's thesis, Ankara University]. Council of Higher Education National Thesis Center.

Ültay, E., Akyurt, H., & Ültay, N. (2021). Sosyal bilimlerde betimsel içerik analizi. *IBAD Sosyal Bilimler Dergisi,* (10), 188-201. https://doi.org/10.21733/ibad.871703

Wei, Y., Subramaniam, G., & Wang, X. (2025) The glass ceiling perception and female teacher burnout: The mediating role of work-family conflict. *Frontiers in Psychology,* (16), 1551903. https://doi.org/10.3389/fpsyg.2025.1551903

Weyer, B. (2007). Twenty years later: Explaining the persistence of the glass ceiling for women leaders. *Women in Management Review, 22*(6), 482-496. https://doi.org/10.1108/09649420710778718

Yamagata, H., Yeh, K. S., Stewman, S., & Dodge, H. (1997). Sex segregation and glass ceilings: A comparative statics model of women's career opportunities in the federal government over a quarter century. *American Journal of Sociology, 103*(3), 566-632. https://doi.org/10.1086/231251

Yelkikalan, N. (2006). Başarılı girişimcilikte cinsiyetin rolü: Kadın ve girişimciler. *Girişimcilik ve Kalkınma Dergisi, 1*(1), 45-54.

Yıldırım, A., & Şimşek H. (2011). *Sosyal bilimlerde nitel araştırma yöntemleri* (8th ed.). Seçkin.

Yıldız, E., & Bedük, A. (2020). Kadın mali müşavirler ve muhasebe çalışanlarında cam tavan sendromunun iş tatminine etkisi. *Toplum Bilimleri Dergisi, 14*(28), 193-221. http://doi.org/10.29228/tbd.2007.29175

Yıldız, S. (2017). Sosyal bilimlerde örnekleme sorunu: Nicel ve nitel paradigmalardan örnekleme kuramına bütüncül bir bakış. *Kesit Akademi Dergisi, 3*(11), 421-442. https://doi.org/10.18020/kesit.1279

Yücedağ, F., & Günbayı, İ. (2016). Eğitim yöneticiliğinden ayrılmış kadınların yaşadığı sorunlara ilişkin görüşleri: Fenomenolojik bir araştırma. *Eğitim ve Öğretim Araştırmaları Dergisi, 5*(4), 86-99.

Education and Science

Original Article

Vol 51 (2026) No 226 481-514

Eighteen years of the Journal of Education and Science (2007-2024): publication trends, research gaps, and global representation

Eray Selçuk [1], Alexander Amigud [2], Hasan Özcan [3], Abdulkadir Özkaya [4]

Abstract

Education and Science (Eğitim ve Bilim) is a leading peer-reviewed educational research journal published by the Turkish Education Association. To examine research trends, we conducted a bibliometric analysis and topic modeling on a dataset of n=1,372 articles published in the journal from 2007 to 2024 using the PRISMA approach. Our analysis identified n=2,077 unique authors, with n=421 (20%) publishing either as sole authors or having contributed only once. The co-authorship rate is 2.06%, and the international collaboration rate is 3.72%. The journal's annual growth rate of publications has declined by 2.73%, with the mean article age of 8.77 years and 5.39 citations per article. The analysis further identified n=4,088 unique keywords and n=61349 cited sources, which were thematically organized into eight research areas as follows: [1] Education and Management in Higher Education (3.92%), [2] Innovative Approaches in Teacher Education and Teaching Skills (6.24%), [3] Language Teaching and Learning Approaches (6.40%), [4] Psychosocial Dynamics and Organizational Structures in Educational Environments (7.49%), [5] Inclusion in Education and Education Policies (13.53%), [6] Enhancing Early Childhood Education and Skills (18.39%), [7] Interdisciplinary Pedagogical Approaches (20.72%), and [8] The Effects of Psychological Factors on Student Success in Education (23.31%). Moreover, we examined the institutional and national affiliations of researchers whose studies were published in the journal. The findings highlight that Hacettepe University in Turkey had the highest number of contributing authors. As well, an analysis of the keywords used in these articles, revealed that most studies centered on teaching and learning. We discuss the trends observed in the number of citations over the 18-year period and conclude with recommendations for fostering the journal's growth.

Keywords

Education and Science
Bibliometric analysis
Text mining
Topic modeling
Big data
Machine learning
Publication trends

Article Info

Received: 12.18.2024
Accepted: 06.30.2025
Published Online: 07.23.2025

DOI: 10.15390/ES.2025.2535

[1] Ministry of National Education, Ankara, Türkiye, erayselcuk84@gmail.com
[2] International Business University, Faculty of Management, Toronto, Canada, aamigud@ibu.ca
[3] Aksaray University, Faculty of Education, Department of Science Education, Aksaray, Türkiye, hozcan@aksaray.edu.tr
[4] Hatay Mustafa Kemal University, Faculty of Education, Department of Science Education, Hatay, Türkiye, aozkaya@mku.edu.tr

Introduction

Educational research is a continuously evolving field that strives to meet the changing needs of learners, instructors, and society. Ferguson et al. (2017) have argued that the role of education is threefold, encompassing personal growth, civic responsibility, and economic progress. These aspects are reflected in education's aims of enhancing employment and labor efficiency, developing active citizens capable of addressing global challenges, and fostering personal growth. To achieve these goals, it is important to recognize that the education system is a complex mechanism that requires a thorough understanding and sound data for evaluation. This is why educational research plays a pivotal role in educational practices. The review and systematization of findings are timely and important due to the continuous nature of research.

In their seminal work titled Educational Research: A Guide to the Process, Wallen and Fraenkel (2013) highlight two core goals that educational research serves. They stress that "a major purpose of educational research is to provide evidence to help people decide which opinions are correct—or at least more correct…A second purpose of educational research is to help us develop better ways to think about the field of education" (p.3). They further noted that" [d]ecisions affecting the greatest resource that human beings have, their children, should be informed by knowledge, not by the loudest voices that can be heard" (p.3).

This warrants an examination of educational research to assess the state of the art, identify key topics and thought leaders, and highlight emerging trends. Conducting a search in the "Education & Educational Research" category on Web of Science will yield thousands of results (Huang et al., 2020). Systematic reviews of academic journals are indispensable for examining the state of the art in educational research and advancing the field by identifying gaps in knowledge and charting a path forward for future inquiries. The insights gained from systematic reviews are valuable not only to practitioners but also to editorial boards and publishers.

Education and Science Journal (Eğitim ve Bilim), published by the Turkish Education Association (TED) and indexed in the Web of Science (WoS) Q4, is a leading peer-reviewed educational research journal that has been published quarterly since 1976 in both English and Turkish.

Its editorial board is composed of academics from international and national universities. It is Turkey's only journal in the field of educational sciences indexed in the SSCI of WoS. Its 2023 Journal Citation Indicator is 0.26 (Szomszor, 2024), and its H5 index is 24 (Google Scholar, 2024). The journal brings together researchers who strive to solve educational problems at both local and international levels. Considering its scope, multidisciplinary focus, and notable influence on educational research, both within Türkiye and internationally, Education and Science (Eğitim ve Bilim) is an important voice in the educational research community, and therefore, was chosen for analysis. While it is important to identify the trends of this journal, which aims to contribute to high-quality scientific studies addressing issues in learning and teaching and aspires to become a "focus journal" as stated in its mission statement, these objectives provide the impetus for this study.

This study aims to reveal current trends, citation patterns, and research collaborations by analyzing publications in the field of educational sciences. Understanding these trends is imperative for promoting scientific productivity and guiding future research. Furthermore, this analysis will help assess the impact of academic journals and provide strategic recommendations for the development of educational research.

This research is important and timely because through the lens of the journal, we can gain valuable insights into the state of the art that can enhance our understanding, better inform educational practices and student experiences, and guide future directions in educational research.

Study aims and research questions

The study aims to answer the following four questions:

1. Who are the researchers that publish in the journal?
2. What topics and keywords emerge most frequently in the studies?
3. Which institutions and countries are represented by researchers in studies published in the Education and Science Journal?
4. What trends are observed in the number of citations over the 18-year period (2007-2024)?

Background

Much of the literature on educational research employs bibliometric analysis methods to examine large corpora of academic texts spanning several decades. As will be evidenced in the following paragraphs, the results underscore the breadth and depth of research interests and illustrate how educational research has transformed over time. The term bibliometrics is a portmanteau term combining two words— "bibliography" and "metrics," coined by Alan Pritchard in his paper titled Statistical Bibliography or Bibliometrics?, published in the Journal of Documentation (Agada, 1987). Pritchard defined the term as "the application of mathematics and statistical methods to books and other media of communication" (Pritchard, 1969 as cited in Roemer & Borchardt, 2015). More broadly, bibliometrics serves to quantify and analyze textual content.

A recent example of bibliometric analysis is the study published in Educational Review by Huang et al. (2020), which examined a sample of n=19,084 papers published between 2000 and 2017. The results highlighted a sixfold increase in publication volume over this period. The most frequently researched topics were higher education, interactive learning environments, and teaching and learning strategies. The USA, UK, and Australia produced the highest number of publications annually.

Another study by Hallinger and Kovačević (2019), examined n=22,361 peer-reviewed articles on educational administration published in Scopus-indexed journals between 1960 and 2018. The study employed a science mapping technique to analyze research trends in the field, identifying influential papers and the institutions actively engaged in this area of research. The study denotes a shift in focus of research over time from "school administration" to "school leadership".

An example of a study specific to the Education and Science journal is that by Selçuk et al. (2014) which analyzed the content of n=492 articles published between 2007 and 2013. The study examined variables such as sample size, method, discipline, subject, data collection tools, and summarized core findings and recommendations. The results suggest that much of the discourse revolved around curriculum and instruction. The studies often relied on quantitative methods, had small to medium sample sizes, collected data through scales, compared means of variables, and were primarily authored by scientists from Hacettepe University.

A more recent bibliometric analysis by Altunışık (2023) examined a sample of n=1,270 articles published in the Education and Science Journal covering the period from 2007 to 2021 analyzing the total number of publications and citations, h-index, citations per publication, annual citation rate, citation threshold, and total link strength. The results confirm those of Selçuk et al. (2014), highlighting that much of the research is published by scholars from Hacettepe University, which was cited n=5,361 times. The most cited publication was in the field of science education by Murat Özdemir. Moreover, the study identified the most frequent keywords; these include: Academic achievement, reliability, validity, teacher, gender, pre-service teachers, self-efficacy, higher education, and structural equation modeling.

Hacettepe University was again mentioned in the study by Gülmez et al. (2021), who conducted a bibliometric analysis of educational research in Turkey published in international journals, using a sample of n=613 articles. The results highlight that much of the research was authored by researchers from Hacettepe University, particularly in the fields of educational technologies and science education. Moreover, articles published by Middle East Technical University received the most citations.

While the scope of these studies was quite broad, some scholars take a more specific focus, such as the study by Dao et al. (2023), which examined a sample of 309 papers published during the 2017-2021 period on Education 4.0 an emerging field related to educational transformation in the context of the Fourth Industrial Revolution. In line with the bibliometric genre, the authors identified countries with the highest number of publications, where Malaysia, Indonesia, Mexico, Brazil, and the Philippines are among the top contributors, while international collaborations are scarce.

While the studies mentioned above are methodologically similar to this one, the following studies are comparable in terms of content.

A study with a more specific focus on educational research from Turkey was conducted by Özcan and Akar (2024) who analyzed a modest sample of n=176 articles published in the National Journal of Education between June 2021 and January 2024 using bibliometric analysis. These results suggest that much of the research was conducted using qualitative methods, revolved around the topics of educational sciences and management, and that most articles were authored by teachers from the Ministry of National Education (MoNE).

A more recent study by Rojas- Rojas-Sánchez et al. (2023) examined the application of virtual reality technologies in education, using bibliometric analysis to identify key research topics, influential authors, sources, and citation trends across a sample of n=273 articles published between 2010 and 2021. The findings suggest that the highest average citation rates were in 2010 and 2014, with 6.9 and 7.3 citations per year, respectively. From 2010 to 2017, research focused primarily on using computer simulations to enhance student engagement. In 2018, the focus shifted toward social learning, with virtual communities like “Second Life,” where users interact through avatars and communicate via text and voice messaging. In 2019, the theme of designing technology-enhanced learning environments emerged, and by 2020, e-learning became central to the discourse, with increased attention to virtual and augmented reality tools supporting remote learning.

These results tally with those of Johri et al. (2024) who examined a sample of 66 studies published between 2012 and 2023 on the role of the Metaverse in sustainable development to identify emerging themes, co-authorship patterns by country, and influential theories. The results highlight an annual increase in publication volume of 37% and a notable spike in interest in the topic during the COVID-19 pandemic due to the transition to digital communication.

Bibliometric analysis using topic modeling-based methods is scarce in literature. A search of the literature revealed that only one study by Özyurt and Ayaz (2022) was found, in which a complete journal review was conducted using the aforementioned method. Their study was conducted to mark the 25th anniversary of the Education and Information Technologies (EAIT) journal. Topic modeling-based bibliometric analysis was employed, and the results were organized into 21 topics. The top five most frequently discussed topics include “Technology acceptance”, “Social network-based learning”, “Teacher education”, “E-learning satisfaction” and “E-learning”. The study further identified two topics that are growing in interest; these include: gamification and technology acceptance.

To mark the 50-year anniversary of The British Journal of Educational Technology (BJET), Chen et al. (2020) conducted a bibliometric analysis and topic modeling of n=3,710 articles published between 1971 and 2018 in the WoS database. The authors identified publication and citation trends, examined the distribution of article types, discovered the most relevant countries, affiliations, and authors, and uncovered key thematic features by analyzing publication abstracts and titles using word cloud and topic modeling analysis. Technology-supported classroom pedagogy, blended learning, online social communities, mobile-supported language learning, game-based learning, and socialized e-learning were identified as topics of increasing relevance to educational needs.

In another study, Yun (2020) used topic modeling to determine research trends in physics education by analyzing 2,959 articles from the American Journal of Physics (AJP) and n=745 articles from Physics Review Physics Education Research (PRPER). A total of 13 topics were extracted from the two journals, where "pedagogical knowledge content", "assessment of achievement" and student's "gender" were identified as topics of growing interest. And while "teacher education" and "students' reasoning process" were identified as topics of high interest, the results suggest that attention to "introduction to physics" and "problem solving" is declining.

Given the scope of prior research, literature reviews in the field of education are generally limited to bibliometric analyses, and subject modeling was found to have been conducted in education-related journals in only three studies. While bibliometric studies provide an overview of topics or literature within a specific time period, topic modeling allows for a more in-depth examination. In this study, the Journal of Education and Science was examined not only through a bibliometric lens but also through topic modeling, to provide a holistic view of the potential research gaps and future publication trends.

Method

To examine research trends in the field of educational research, this study employs bibliometric analysis (Donthu et al., 2021) and topic modeling techniques (Vayansky & Kumar, 2020). These approaches are increasingly used for the systematic evaluation of academic literature, as advances in information and communication technologies make it feasible to analyze large, thematically organized academic corpora that are accessible and affordable through databases. Donthu et al. (2021) maintain that "bibliometric analysis is useful for deciphering and mapping the cumulative scientific knowledge and evolutionary nuances of well-established fields by making sense of large volumes of unstructured data in rigorous ways" (p. 285).

Systematic Review

The bibliometric profiles of the studies in the Education and Science Journal were examined and the topics under which the studies were grouped were determined. This type of research is classified as descriptive research. Descriptive research is research that tries to describe the existing situation in order to provide insights to generate hypotheses and theoretical foundations (Aggarwal & Ranganathan, 2019; Erkuş, 2021). By observing and collecting data on a particular topic, descriptive research helps researchers gain a deeper understanding of a particular topic and provides valuable insights that can inform future studies. This method is applied to large samples, especially in quantitative research. The researcher collects data from such surveys and analyzes it to better understand the current situation. Descriptive models often help researchers comprehend the existing body of knowledge, identify research questions and guide the hypothesis development process before embarking on a research project (Köroğlu, 2015).

Data Collection

The data of the study was obtained by searching the Education and Science Journal from the Web of Science (WoS) database. The Web of Science database is a database developed by Clarivate Analytics and contains more than a thousand journals, books, conferences and reports. This database includes the Science Citation Index-Expanded (SCI-Expanded), Social Sciences Citation Index (SSCI), Arts & Humanities Citation Index (A&HCI), Emerging Sources Citation Index (ESCI), Book Citation Index (BCI) and Conference Proceedings Citation Index (CPCI) indexes (Li et al., 2018). Education and Science Journal is included in the Q4 quarterly journal list in the SSCI index. While collecting the data, the following key index was used:

"SO= (EGITIM VE BILIM EDUCATION "AND" SCIENCE)"

Inclusion and exclusion criteria were applied to the data obtained as a result of the screening according to the PRISMA flowchart created by Moher et al. (2009) The PRISMA flow used in the process of organizing the data of the study is given in Figure 1.

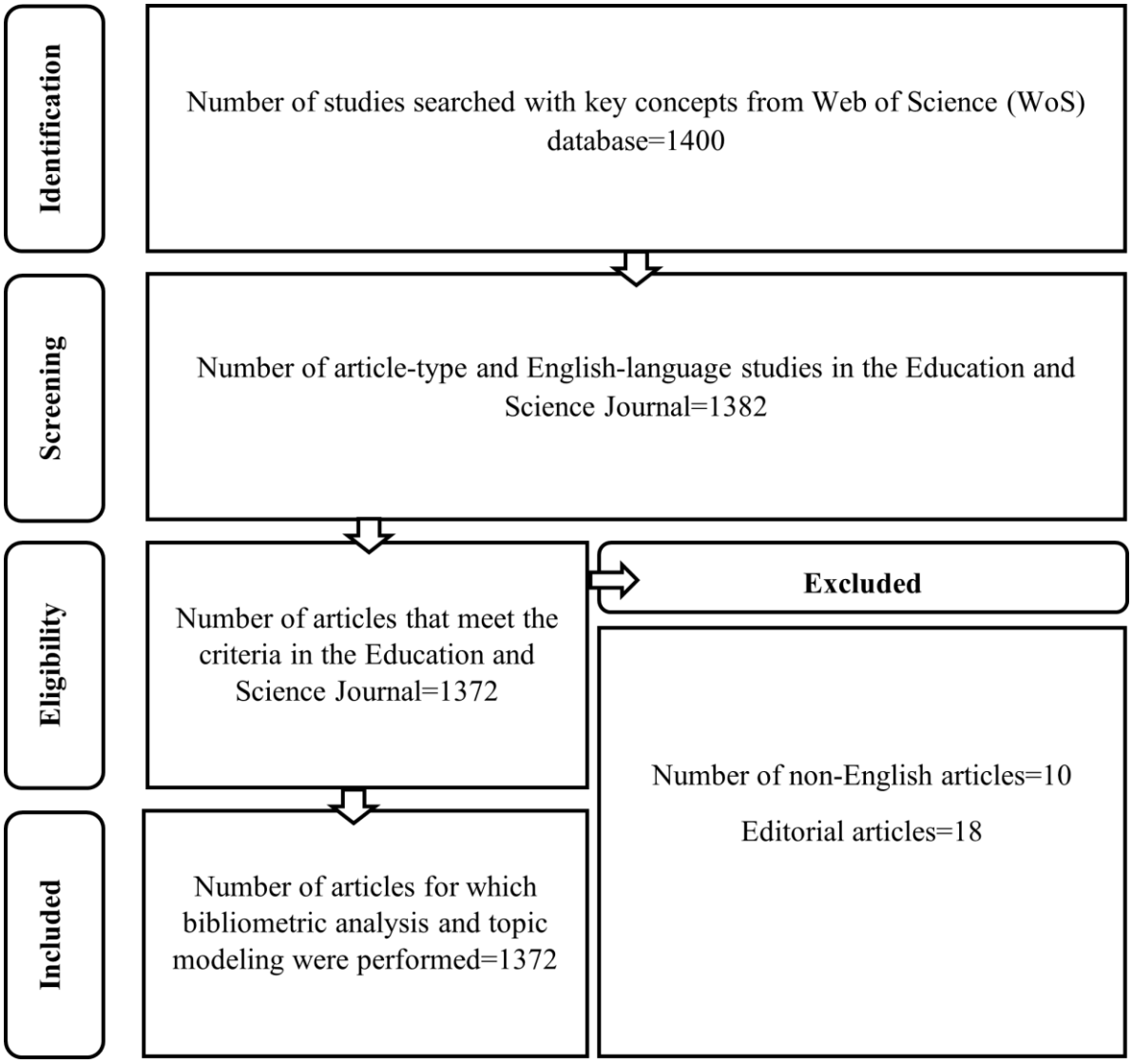


Figure 1. The process of collecting the data according to the PRISMA flow diagram

When Figure 1 is examined, n=1,372 publications published in the WoS database between 01/01/2007 and 31/07/2024 in the Education and Science Journal in the article type and in English constitute the data of the study.

Data Analysis

Bibliometric analysis is a tool for understanding the structure, transformation, and impact of scientific literature. It helps to understand relationships within and assess the impact of scientific research (Zupic & Cater, 2015). Bibliometric analysis was performed with the Bibliometrix package written by Aria and Cuccurullo (2017) in the R programming language. The analysis was visualized with the Biblioshiny application, which is a user-friendly interface within this package. As a result of the analysis, the distribution of articles by year, the author who produced the most articles in the journal, the most cited author, the most cited article, the most frequently used keywords in the articles and trending topics were reported.

In the second step of the data analysis, topic modeling analysis was performed. Topic modeling is a text mining method. Text mining is one of the challenging research topics due to the need to organize and categorize an increasing number of electronic documents worldwide (Uysal & Gunal, 2014). The goal of text mining is to classify, group, and tag texts; summarize data sets, create taxonomies, and gain knowledge by analyzing word frequencies and relationships between data items (Rouse, 2018). Text mining is an increasingly popular method for extracting important knowledge and insights from large volumes of unstructured textual data. As the number of research articles, reports, and other text-based resources in higher education continues to grow, text mining provides a powerful method to analyze and understand research trends in this field (Tyagi, 2021). Topic modeling is a text mining technique that aims to find hidden themes and patterns in a text collection. By eliminating the need for manual categorization or tagging, this unsupervised machine learning approach gives researchers the power to grasp the organization and content of extensive document collections. It independently distinguishes hidden motifs and groups of related terms or concepts embedded in textual data, thus facilitating a comprehensive understanding of the content (Shadrova, 2021). Topic modeling is the process of identifying important topics in a text using probabilistic models. These probabilistic models help to analyze and classify a corpus of texts so that a person can easily recognize important topics in them without having to read the entire document. It is important to note that these models are informed by knowledge about the texts being analyzed, the topics of interest can range from science to art. Models that fall into this category include latent semantic analysis, probabilistic latent semantic analysis and Latent Dirichlet Allocation (LDA) (Özkan, 2015).

In this study, the LDA method was used in the topic modeling analysis. LDA is a generative probabilistic model. It improves Probabilistic Latent Semantic Discrimination by adding priorson topic distributions for documents and word distributions for topics. LDA assumes that the creation of each document requires an initial sampling of a topic distribution followed by a sampling of words from the identified topics. To estimate model parameters, techniques such as Gibbs sampling or Variational Inference are commonly used and allow for an accurate representation and analysis of the document corpus (Blei et al., 2003).

Text Preprocessing Phase

Before proceeding with topic modeling analysis, the data must undergo a preprocessing process and text-to-vector transformation. Text preprocessing is one of the necessary processes for organizing and structuring unstructured texts prior to analysis (Aggarwal & Zhai, 2013). For this, texts are first broken down into the smallest meaningful units, i.e. words. Then, meaningless and missing words, web links, numerical expressions and punctuation are removed from the texts. After this process, frequently used stopwords in English are deleted from the texts. In this way, the texts are transformed into word vectors. These processes, depicted in Table 1, are as follows (AlSumait et al., 2008; Barde & Bainwad, 2017).

Table 1. Text Preprocessing Steps

#	Process Step	Description
1	Transformation	All words in the corpus are converted to lower case. In addition, removal of diacritics and URL extensions, and html parsing are also performed in this process.
2	Tokenization	All sentences in the corpus are broken down to individual words.
3	Normalization	The original forms of words are extracted. That is, valid root forms are provided, considering the lexicon of a given language.
4	Lemmatization	During lemmatization, the focus is on nouns, adjectives and verbs, ensuring the singularity of words.
5	Filtering	The extraction of stopwords and digits in the corpus that do not make sense on their own is performed

After these steps, word vectors are created, the document term matrix is created, and the topic modeling analysis is started.

Topic Modeling Analysis

The collected texts, initially qualitative data, were transformed through a word-vector transformation, to become suitable for quantitative analysis. At this stage of the analysis, the Orange Data Mining program was used. This program uses the Python language (Demšar et al., 2013). The steps of this process are given in Figure 2.

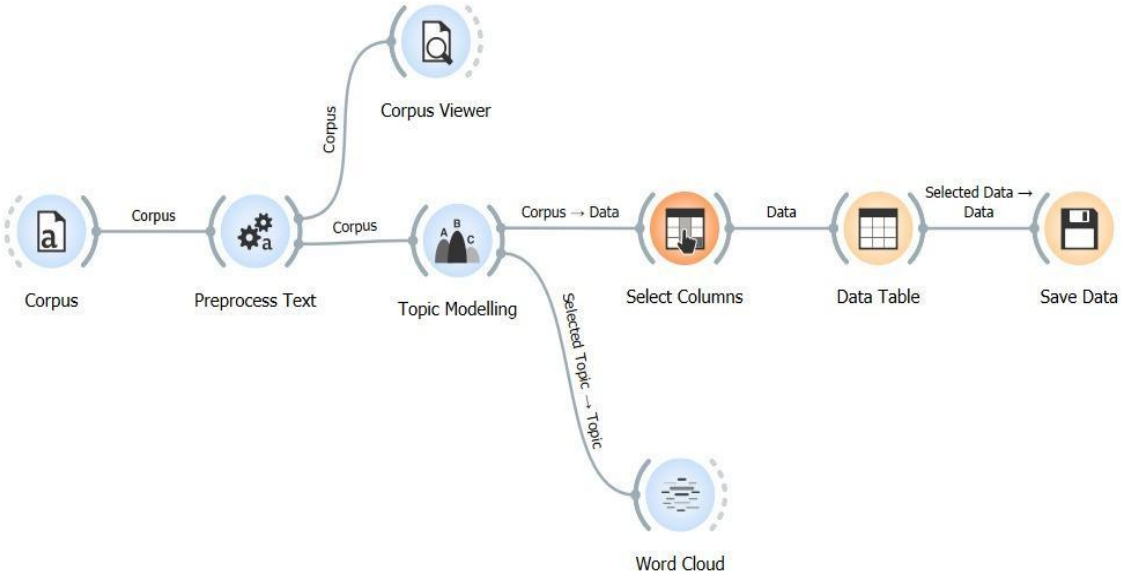


Figure 2. Steps of topic modeling analysis in the Orange Data Mining program

Topic modeling analysis is built on two basic assumptions. The first assumption is that each document in the corpus is a mixture of all topics. The second assumption is that each topic is a mixture of lexical terms. The interpretation that emerges from these two assumptions is that the topic variable in the model is selected repeatedly within each document, allowing the documents to consist of more than one topic. The researcher is at the forefront in determining the number of topics. That is, the decision is subjective and based on expert judgment (Bystrov et al., 2023). In order to account for possible subjectivity and to allow for a more standardized estimation procedure, several evaluation criteria have been developed to determine the optimal number of topics in LDA models. Some of them aim to minimize the similarity of different topics (Cao et al., 2009), maximize topic coherence (Mimno et al., 2011), or maximize the agreement between predicted and actual document term frequencies (Lewis & Grossetti, 2022). Accordingly, one of the methods for calculating the appropriate number of

topics is topic fit analysis. Topic Fit Analysis serves as an evaluation method for topic models designed to determine the ideal number of topics for LDA based content analysis (Alrayashi, 2023). Another method is the calculation of the perplexity value. The perplexity refers to a metric that gives the average uncertainty provided by the model to each word in the dataset (Tang et al., 2018). In general, the lower the model's perplexity score, the better the generalization performance. Before performing the topic modeling analysis, coherence and log perplexity values were calculated to determine the appropriate number of topics. The number of topics increased to 30 starting from 1. The obtained coherence and perplexity values are given in Table 2.

Table 2. Coherence and Log perplexity values of topic modeling analysis

Topic coherence	Log perplexity	Topic count
0.261	130.753	1
0.271	129.483	2
0.300	126.739	3
0.306	125.891	4
0.316	125.944	5
0.314	126.210	6
0.309	125.865	7
0.319	***125.927***	***8****
0.308	126.135	9
0.303	126.424	10
0.314	126.852	11
0.309	126.494	12
0.310	126.297	13
0.302	126.430	14
0.311	126.635	15
0.310	127.440	16
0.306	127.318	17
0.310	127.774	18
0.306	128.332	19
0.307	129.186	20
0.298	129.377	21
0.304	129.678	22
0.298	129.377	23
0.306	130.077	24
0.300	130.895	25
0.310	130.503	26
0.305	131.526	27
0.301	132.119	28
0.310	132.258	29
0.306	132.145	30

* Determined topic count

The highest coherence value (see Table 2) and one of the lowest log perplexity values are obtained when the number of topics is eight. By analyzing both these values and the resulting word clouds, the researchers reached a consensus that the number of topics should be eight.

Trend Analysis

The trends of topics may change over time. This allows researchers to focus on different topics in their field of study. The annual percentage of the number of articles on each topic reflects the degree of interest in that topic. The percentage data changes every year, and some changes fluctuate rapidly in the short term. A scatter plot of the topic's annual share data can reflect this change. We refer to the literature to obtain a relatively stable change trend of the topic based on the annual percentage data (Cho et al., 2017; Yin & Yuan, 2022). Accordingly, in the last part of the analysis, the distribution over time according to the topic weights obtained from the LDA model, the volumes of the topics over time (by years), the change trends of the topics, and the acceleration values of the topics relative to each other were calculated.

Results

In this section, following the order of research questions, we present the results of bibliometric analysis and topic modeling. We commenced the analysis by providing a general overview of the journal, including the number of keywords and references, and information about the authors of n=1,372 articles. Table 3 depicts that the annual growth rate of the Education and Science Journal is -2.73%. annual growth rate is negative. The average age of the articles published in the journal is 8.77 years. This shows that the journal is a young journal depending on the time it has been in the WoS database.

The average number of citations per article is 5.386 where n=4088 keywords and n=61,349 sources were used in the articles. Of these articles, n=421 were written by a single author and the total number of authors was n=2,077.

Table 2. Overview of articles published in the Education and Science Journal

General information		**Results**
Main information about data	Time interval	01/01/2007:31/07/2024
	Number of journals	1
	Number of articles	1,372
	Annual growth rate	%-2.73
	Average age of articles	8.77
	Average number of citations per article	5.386
Articles	Number of keywords	4,088
	Number of references	61349
Authors	Number of authors	2,077
	Number of single authors	421
	Co-authorship rate per article	2.06
	International authorship rate	%3.72

The rate of co-authorship per article is 2.06. Most of the articles have two authors. The international co-authorship rate is 3.72%. Accordingly, it is seen that the rate of foreign authors in the journal is low. The distribution of 1,372 articles in the journal between 2007 and 2024 is presented in Figure 3.

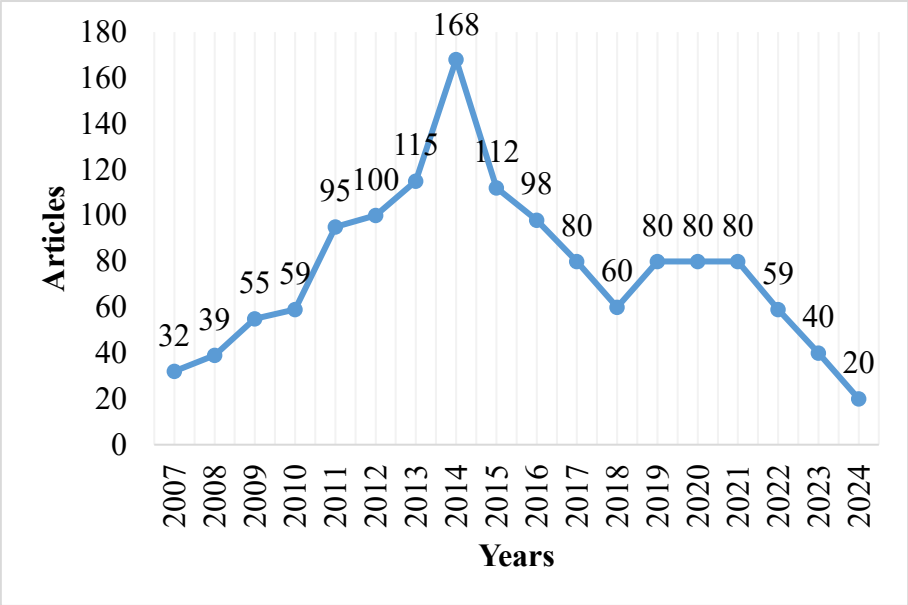


Figure 3. Distribution of articles according to years of publication

The highest number of articles was published in 2014 (n=168), while the lowest number of articles was published in 2024 (n=20). Since 2024 has not been completed yet, it is seen that the least number of articles was published in 2007 (n=32). The yearly average citation rates of articles are depicted in Figure 4.

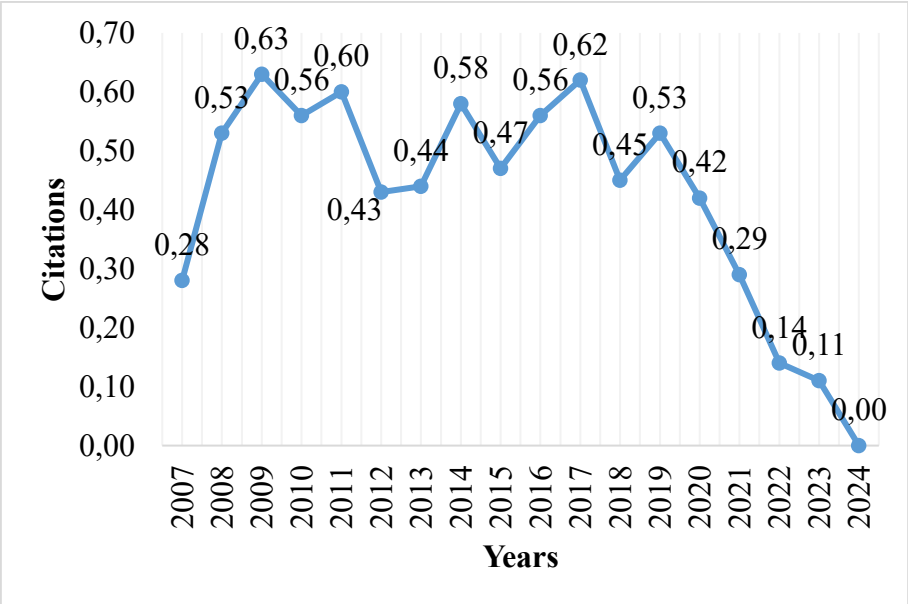


Figure 4. Citation averages of articles by year

The years with the highest average citation rates are 2009 and 2017 (c=0.63). A breakdown of the annual citation rates is presented in Table 4.

Table 3. Annual citation rate

Year	Mean Total Citations per Article	N	Mean Total Citations per Year	Citable years
2007	5.38	32	0.30	18
2008	9.54	39	0.56	17
2009	10.64	55	0.66	16
2010	8.97	59	0.60	15
2011	8.99	95	0.64	14
2012	6.03	100	0.46	13
2013	5.68	115	0.47	12
2014	6.96	168	0.63	11
2015	5.17	112	0.52	10
2016	5.64	98	0.63	9
2017	5.54	80	0.69	8
2018	3.60	60	0.51	7
2019	3.70	80	0.62	6
2020	2.55	80	0.51	5
2021	1.44	80	0.36	4
2022	0.58	59	0.19	3
2023	0.32	40	0.16	2
2024	0.00	20	0.00	1

Although most articles were published in 2014, the highest number of citations did not occur this year. Although 20 articles were published until the first half of 2024, there is no cited publication yet. When the average number of citations per year of the articles was analyzed, it was determined that the ratio was highest in 2017 and lowest in 2024. However, a gradual decrease was observed in the number of articles cited in the journal over the years. While older publications of the journal receive more citations, this number decreases for newer articles. The average number of citations per year between 2008 and 2020 has reached the highest values.

Who are the researchers that publish in the journal?

The number of articles and proportionally decimalized articles of the ten researchers who produced the most articles in the Education and Science Journal are given in Table 5.

Table 4. Ten researchers who produced the most articles in the Education and Science Journal

Authors	Articles	Articles Fractionalized
Nuri Doğan	10	5.00
Murat Özdemir	9	5.39
Ömer Geban	7	3.00
Hayati Akyol	6	2.70
Adnan Baki	6	2.75
Yüksel Göktaş	6	2.08
Ömer Kutlu	6	2.83
Semra Sungur	6	2.67
Ahmet Akın	5	3.50
Uğur Akın	5	2.83

The highest number of articles in the Education and Science Journal was written by researcher Nuri Doğan (n=10). Murat Özdemir (n=9) and Ömer Geban (n=7) are in second and third place respectively (see Table 5). The proportionally decimalized number of articles measures an author's contribution to a set of published articles. Accordingly, the researcher with the highest number of decimalized articles is Murat Özdemir (n%=5.39). Nuri Doğan (n%=5.00) ranked second and Ahmet

Akın (n%=3.50) ranked third. The ten researchers with the highest number of citations among the researchers who published articles in the Education and Science Journal are presented in Table 6.

Table 5. The ten most local cited researchers in the Education and Science Journal

Authors	Total citations
Selahattin Gelbal	15
Duygu Anıl	13
Mehmet Palancı	12
Veli Duyan	11
Cemal Ergin Ekinci	11
Deniz Melanlıoğlu	11
Hakan Dündar	10
Mehmet Kandemir	10
Ziya Selçuk	10
Eren Ceylan	9

The list of ten most cited researchers in the Education and Science Journal is presented Table 6. The three researchers who received the most local citations from the articles published in the Education and Science Journal are Selahattin Gelbal (n=15), Duygu Anıl (n=13) and Mehmet Palancı (n=12). Selahattin Gelbal has 5 articles published in the Education and Science Journal. Four of these articles are ranked first due to their high annual citation rates.

The distribution of the publications of the researchers who published articles in the Education and Science Journal according to the years is given in Figure 5. The points in the figure show the article production nodes. As the node size increases, the number of publications increases. As the node color gets darker, the number of citations increases.

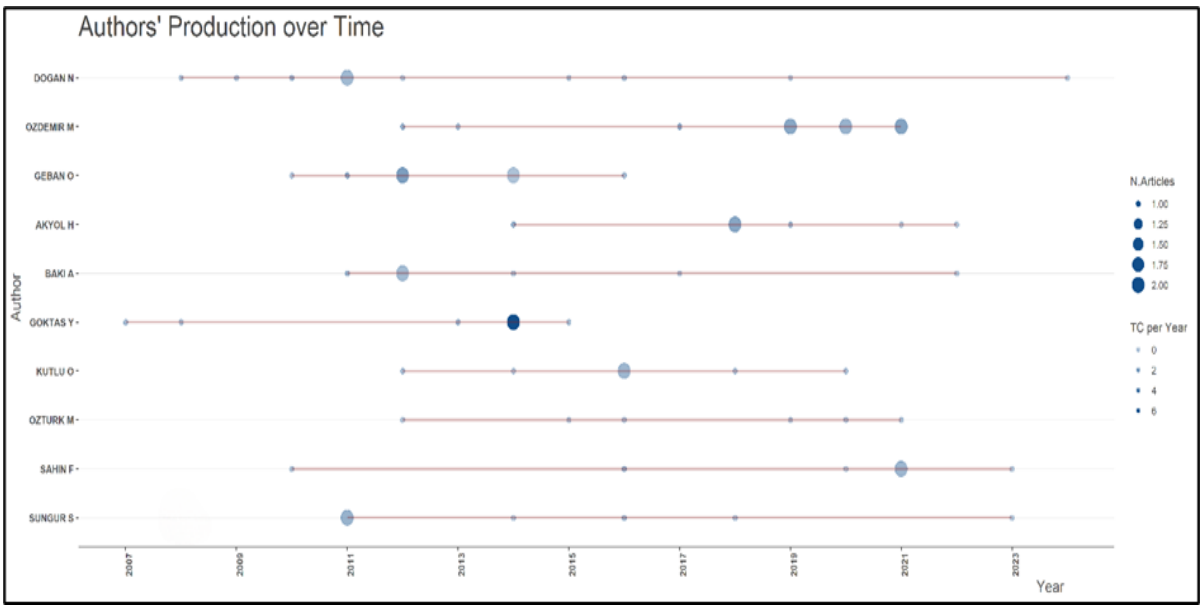


Figure 5. Distribution of the publications of researchers publishing in the Education and Science Journal according to years

Nuri Doğan is the most prolific contributor to the Education and Science Journal, both in terms of the total number of publications and the length of the time span across which these publications were produced. Murat Özdemir and Ömer Geban follow as the second and third most productive authors, respectively. Notably, Murat Özdemir published two articles in each of three consecutive years (2019, 2020, and 2021). Among the researchers with two publications, Yüksel Göktaş stands out with the highest total citation rate per year (see Figure 5).

The publication production rates of researchers publishing in the Education and Science Journal according to Lotka's Law are given in Figure 6. The graph of author productivity according to Lotka's law is a proportional graph showing how many publications authors contributing to a particular field have contributed. Also, this graph allows a quantitative prediction of the authors who will contribute to the relevant literature in the coming years.

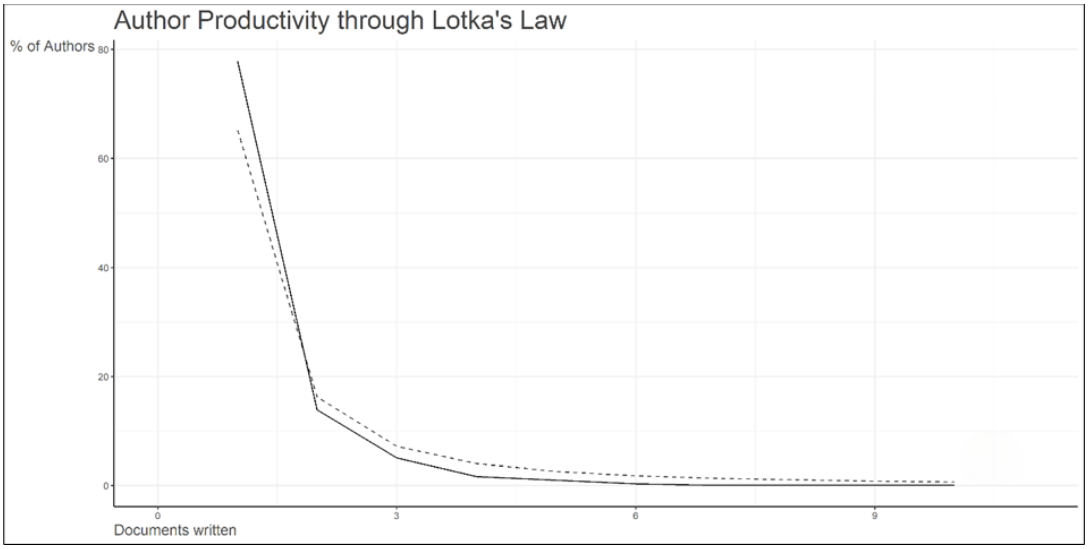


Figure 6. Proportion of researchers publishing articles in the Education and Science Journal producing publications according to Lotka's Law

Out of n=2,077 researchers who published in the Education and Science Journal, n=1,617 published only once. The rate of authors publishing a single article in the journal is approximately 80%, whereas n=289 researchers published two articles. The rate of these researchers is 14%. The remaining researchers published three or more articles (see Figure 6).

What topics and keywords emerge most frequently in the studies?

The ten most frequently used keywords in articles published in the Journal of Education and Science are shown in Figure 7.

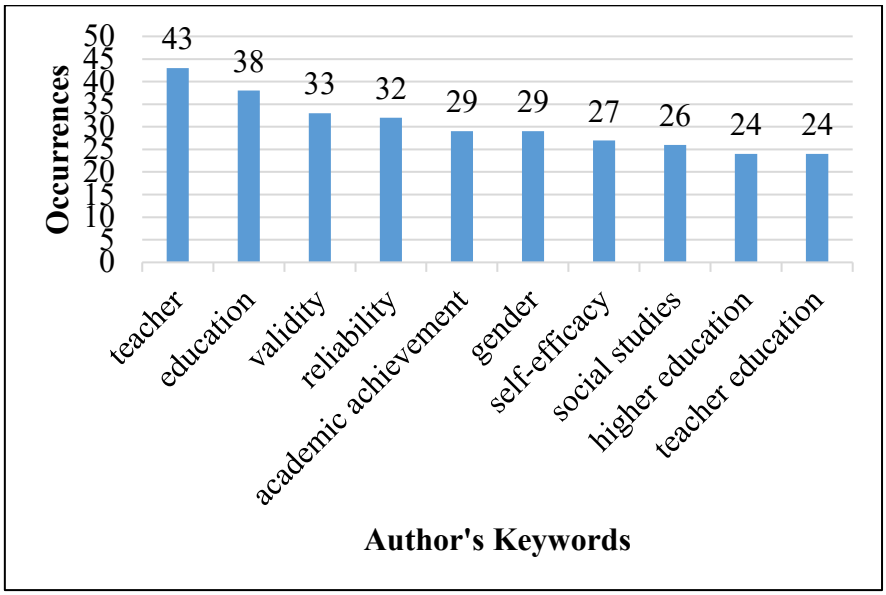


Figure 7. The ten most frequent author's keywords

The top ten most frequently keywords, depicted in Figure 7, include "teacher" (n=43), "education" (n=38), "validity" (n=33), "reliability" (n=32), "academic achievement (n=29)", "gender (n=29)", "self-efficacy (n=27)", "social studies (n=26)", "higher education (n=24)", "teacher education (n=24)". Accordingly, it is seen that most of the articles in the Education and Science Journal are related to the field of teaching. The word cloud consisting of these keywords is given in Figure 8.

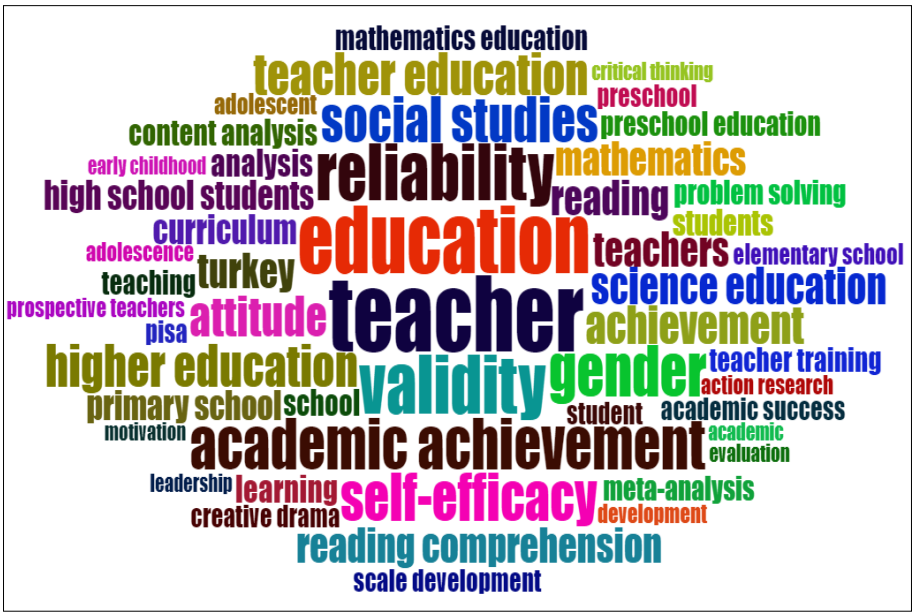


Figure 8. The most frequently used keywords in articles published in the Education and Science Journal

The word cloud in Figure 8 depicts just that. The concepts of "teacher" and "education" are frequently used as keywords in the articles in the journal. The change in the trending topics in the articles published in the Education and Science Journal by year is given in Figure 9.

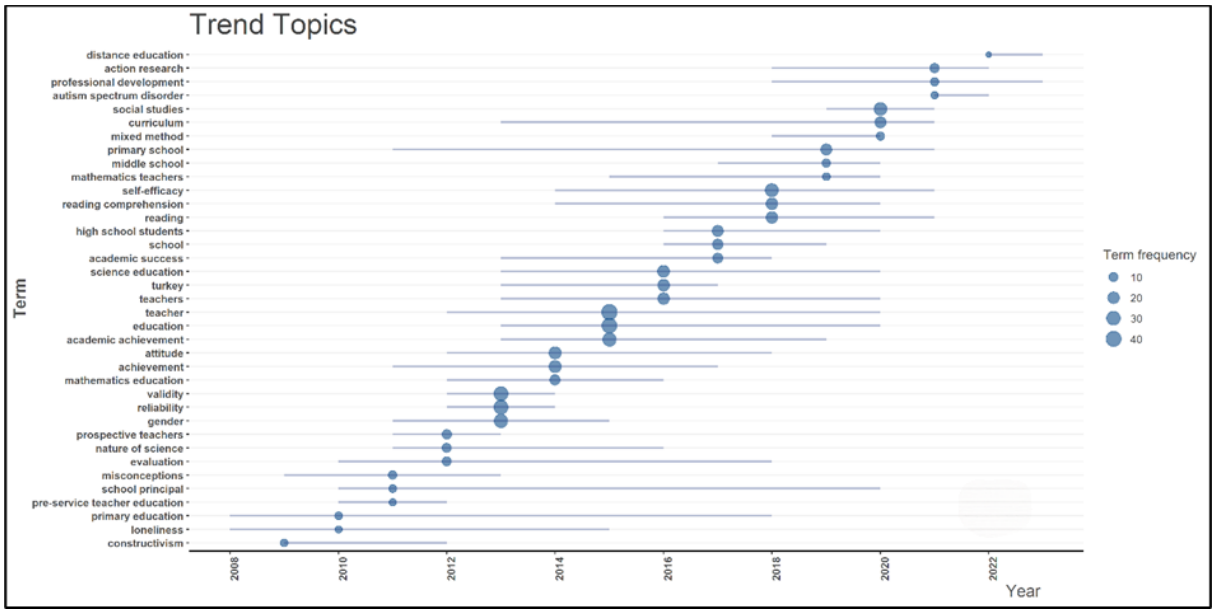


Figure 9. Changes in the trending topics used in articles published in the Education and Science Journal over the years

The temporal distribution of the trending topics used in the articles published in the Education and Science Journal is presented in Figure 9. The first topic used since 2008, when the journal was indexed in the WoS database, is the concept of "constructivism", which was frequently mentioned in education programs in those years. In recent years, the trending topical term due to the effects of the pandemic on the education and training process is the concept of "distance education". The most frequently used topics are the same as the keywords mentioned earlier. These are "teacher", "education", "academic achievement", "validity", "reliability", and "gender". These topics were found in 40 or more articles. According to the years, the topics with the longest lasting impact on the published articles are curriculum, primary school, school principal, and primary education. These topical areas show the dominant topics in the articles published in the Education and Science Journal.

Which institutions and countries are represented by researchers in articles published in the Education and Science Journal?

The top ten academic institutions and the number of articles published in the Education and Science Journal are presented in Table 7.

Table 6. Academic institutions and number of articles published in the Education and Science Journal

Authors' Institutional Affiliations	Articles
Hacettepe University	341
Ankara University	238
Gazi University	228
Anadolu University	109
Marmara University	75
Sakarya University	66
Abant Izzet Baysal University	59
Necmettin Erbakan University	58
Karadeniz Teknik University	56
Dokuz Eylül University	52

Table 7 depicts academic institutions and the number of articles published by their researchers. The highest number of articles in the Education and Science Journal is from Hacettepe University (n=341). Ankara University (n=238) ranked second, and Gazi University (n=228) ranked third. Because the contributions of these three universities to the field of education and their faculties of education have a long history, it is expected that these three universities come first. Many prominent educational scientists work in these universities. The distribution and co-authorship rates of the principal investigators who published articles in the Education and Science Journal according to their countries are given in Table 8.

Table 7. Distribution and co-authorship rates of articles published in the Education and Science Journal according to the affiliation countries of the principal investigators*

Country	Articles	Rate (%)	SCP**	MCP***
Türkiye	1,233	89.9	1,202	31
TRNC	13	0.9	10	3
USA	7	0.5	3	4
Canada	4	0.3	2	2
China	4	0.3	2	2
Malaysia	4	0.3	3	1
Serbia	3	0.2	2	1
Slovenia	3	0.2	3	0
Iran	2	0.1	1	1
Montenegro	2	0.1	1	1

*The table includes only the top ten countries. **Single Country Publications. ***Multiple Country Publications

The researchers responsible for publishing articles in the Education and Science Journal are mostly from Turkey (see Table 8). Almost 90% of the articles were written by researchers from Turkey. Accordingly, it is determined that Turkey is the dominant country in the journal. This is an expected situation since the journal is a publication originating from Turkey. Thirty-one of these articles were made with co-researchers from different countries. Turkey is followed by the TRNC with n=13 articles and the USA with n=7 articles. There are n=4 articles from Canada, China and Malaysia, n=3 articles from Serbia and Slovenia, and n=2 articles from Iran and Montenegro. The inter-country collaboration network of articles published in the Education and Science Journal is given in Figure 10.

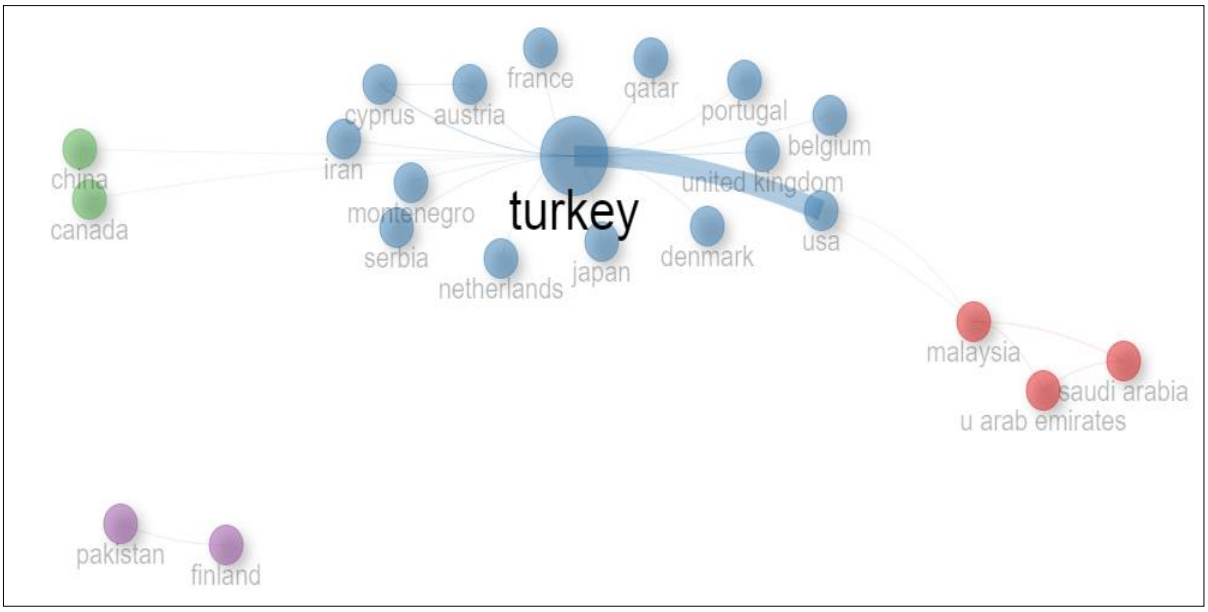


Figure 10. Cross-country collaboration network for articles published in the Education and Science Journal

Most of the articles published in the Education and Science Journal are jointly published between Turkey and the USA (see Figure 10). Turkey has produced articles with 16 other countries. These countries are TRNC, Iran, Montenegro, Serbia, the Netherlands, Denmark, Japan, the United Kingdom, Belgium, Portugal, Qatar, France, Austria, Canada, China, and Malaysia. There are also articles co-published by China and Canada, Pakistan, Finland, Malaysia, Saudi Arabia and the United Arab Emirates. This highlights the journal's commitment to international engagement in its publication practices.

Topic Modeling Using Latent Dirichlet Allocation

The key topics in the articles published in the Education and Science Journal and the changes of topics over time were examined through topic modeling. Based on coherence and log perplexity values, the researchers reached a consensus on eight main topics. The subject terms that were a factor in determining these topics are given in Table 9.

Table 8. Main topics, topic terms and topic weighting ratios resulting from topic modeling

#	Topic name	Topic terms	Rate (%)
1	Education and Management in Higher Education	Study, use, education, universities, teaching, analysis, student, teacher, development, relationship	3.92
2	Innovative Approaches in Teacher Education and Teaching Skills	Teacher, education, student, working, using, teaching, self, service, skills, data	6.24
3	Language Teaching and Learning Approaches	Student, learning, study, school, analysis, research, approach, impact, language, use	6.40
4	Psycho-social Dynamics and Organizational Structures in the Educational Environment	Teacher, study, use, school, student, scale, analysis, data, research, results	7.49
5	Inclusion in Education and Education Policies	Education, teacher, study, student, school, use, research, data, child, inclusion	13.53
6	Enhancing Early Childhood Education and Skills	Study, reading, child, group, student, test, use, skills, teacher, education	18.39
7	Interdisciplinary Pedagogical Approaches	Student, study, teacher, use, research, group, science, data, school, analysis	20.72
8	The Effects of Psychological Factors on Student Success in Education	School, student, teacher, study, self, use, scale, analysis, data, success	23.31

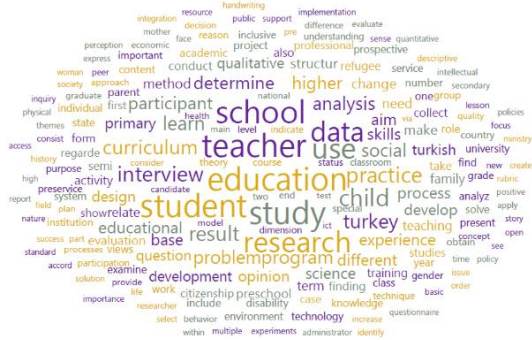
The topics presented in Table 9, constitute the focus of the topic terms. There are eight different main topics created according to the topic terms. However, the weight ratios of the topics indicate the relative importance of the articles published in the Education and Science Journal. Accordingly, "The Effect of Psychological Factors in Education on Student Achievement" has the highest weight rate. The main focus of the Education and Science Journal is the articles within the scope of this topic. One of the most frequently used keywords in the bibliometric findings of the study is "academic achievement". These two findings are consistent with each other. The second most important topic was "Interdisciplinary Pedagogical Approaches" and the third was "Early Childhood Education and Enhancing Skills". "Innovative Approaches in Teacher Education and Teaching Skills" is the last topic in the relative importance order of the journal. In the light of these findings, it can be said that in the relative importance order of the articles published in the Education and Science Journal, student academic achievement emerges first, followed by studies related to teaching skills. The word clouds that are effective in the formation of these main topics are given in Figure 11.

Topic 1: Education and Management in Higher Education

Topic 2: Innovative Approaches in Teacher Education and Teaching Skills

Topic 3: Language Teaching and Learning Approaches

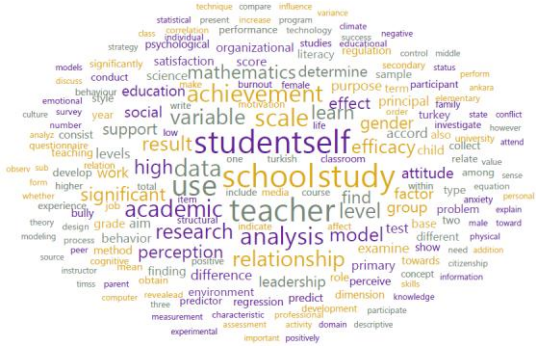
Topic 4: Psycho-social Dynamics and Organizational Structures in the Educational Environment



Topic 5: Inclusion in Education and Education Policies

Topic 6: Enhancing Early Childhood Education and Skills

Topic 7: Interdisciplinary Pedagogical Approaches



Topic 8: The Effects of Psychological Factors on Student Success in Education

Figure 11. Word clouds for main topics

Figure 11 depicts the keywords that define the composition of topic headings. The relative size and position of the keywords are determined by their frequency in the word cloud; the topic is named

based on the ten most frequent keywords. Five examples of articles with the highest weight in the relevant main topics are presented in Table 10. These articles can be accepted as evidence for the validity of the topics formed as a result of topic modeling.

Table 9. The imprint information of the articles published in the Education and Science Journal and collected under the main topics

Topic names	Articles' imprints
Education and Management in Higher Education	Dimici et al. (2016), Selekler-Gökşen et al. (2016), Uslu (2016), Üstünlüoğlu (2016), Zhou and Wu (2016)
Innovative Approaches in Teacher Education and Teaching Skills	Demirtaş et al. (2011), Erten (2022), Gündoğdu (2011), Temli et al. (2013), Ursavaş et al. (2014)
Language Teaching and Learning Approaches	Bican and Demir (2018), Indriyani et al. (2023), Genç (2014), Karababa and Karagül (2013), Mirici et al. (2010)
Psycho-social Dynamics and Organizational Structures in the Educational Environment	Ağaoğlu et al. (2012), Akman and Özdemir (2018), Buyruk and Akbaş (2021), Büyükşahin Çevik (2017), Kılınç (2014)
Inclusion in Education and Education Policies	Aslan (2021), Demir Başaran (2020), Gültekin Toroslu (2013), Şirin (2010), Zayimoğlu Öztürk (2021)
Enhancing Early Childhood Education and Skills	Gür et al. (2017), Özgün et al. (2020), Polat (2021), Tozduman Yaralı and Güngör Aytar (2020), Ünal and Aral (2014)
Interdisciplinary Pedagogical Approaches	Baran and Ata (2013), Kavacık et al. (2015), İlhan and Oruç (2019), Susar Kırmızı (2015), Uzun et al. (2022)
The Effects of Psychological Factors on Student Success in Education	Acar Güvendir (2014), Arslan (2016), Atik and Özer (2020), Hotaman and Yüksel-Şahin (2010), Ünsal Özberk et al. (2018)

Table 10 shows the topics under which the articles published in the Journal of Education and Science are grouped. These are the articles with the highest weight in the relevant main topics. At the same time, these articles can be accepted as evidence of the validity of the topics formed as a result of topic modeling.

What trends are observed in the number of citations over the 18-year period (2007-2024)?

The results suggest that there is no notable change in the distribution of articles across the topics. While there is an increase in the number of articles in some topics in certain years, there is a decrease in some years. However, when the total number of articles produced in the topics are analyzed, it is concluded that the highest number of articles are on "The Effects of Psychological Factors in Education on Student Success". This is followed by "Interdisciplinary Pedagogical Approaches" and "Early Childhood Education and Increasing Skills". The fewest number of articles in the journal are on "Education and Management in Higher Education". The distribution of the topics formed by the articles published in the Education and Science Journal over time is presented in Table 11.

Table 10. Distribution of the topics of the articles published in the Education and Science Journal

Topics	Years																		
	2007	**2008**	**2009**	**2010**	**2011**	**2012**	**2013**	**2014**	**2015**	**2016**	**2017**	**2018**	**2019**	**2020**	**2021**	**2022**	**2023**	**2024**	**Total**
Innovative Approaches in Teacher Education and Teaching Skills	3	2	6	7	7	5	9	15	8	9	4	0	4	2	4	2	2	1	90
Psycho-social Dynamics and Organizational Structures in the Educational Environment	0	1	2	3	5	10	12	16	6	6	7	5	9	2	8	6	3	0	101
Inclusion in Education and Education Policies	8	6	10	7	12	10	20	12	19	8	10	5	8	6	17	9	9	4	180
Interdisciplinary Pedagogical Approaches	12	13	12	13	15	17	19	22	29	19	18	10	17	27	16	14	10	6	289
The Effects of Psychological Factors on Student Success in Education	4	12	13	13	31	26	21	45	21	26	13	17	15	18	15	15	9	1	315
Education and Management in Higher Education	0	0	0	2	2	5	5	5	5	12	3	2	2	4	2	3	1	0	53
Enhancing Early Childhood Education and Skills	5	1	6	10	20	21	25	37	14	14	20	14	18	17	15	10	4	8	259
Language Teaching and Learning Approaches	0	4	6	4	3	6	4	16	10	4	5	7	7	4	3	0	2	0	85
Total	32	39	55	59	95	100	115	168	112	98	80	60	80	80	80	59	40	20	1372

The results suggest that the Education and Science Journal gives more importance to articles on increasing the academic achievement of students and different methodological approaches that serve this purpose, as well as articles related to early childhood, which are thought to affect academic achievement. The volumetric changes and trends of the topics of articles published in the Education and Science Journal over time are depicted in Figure 12.

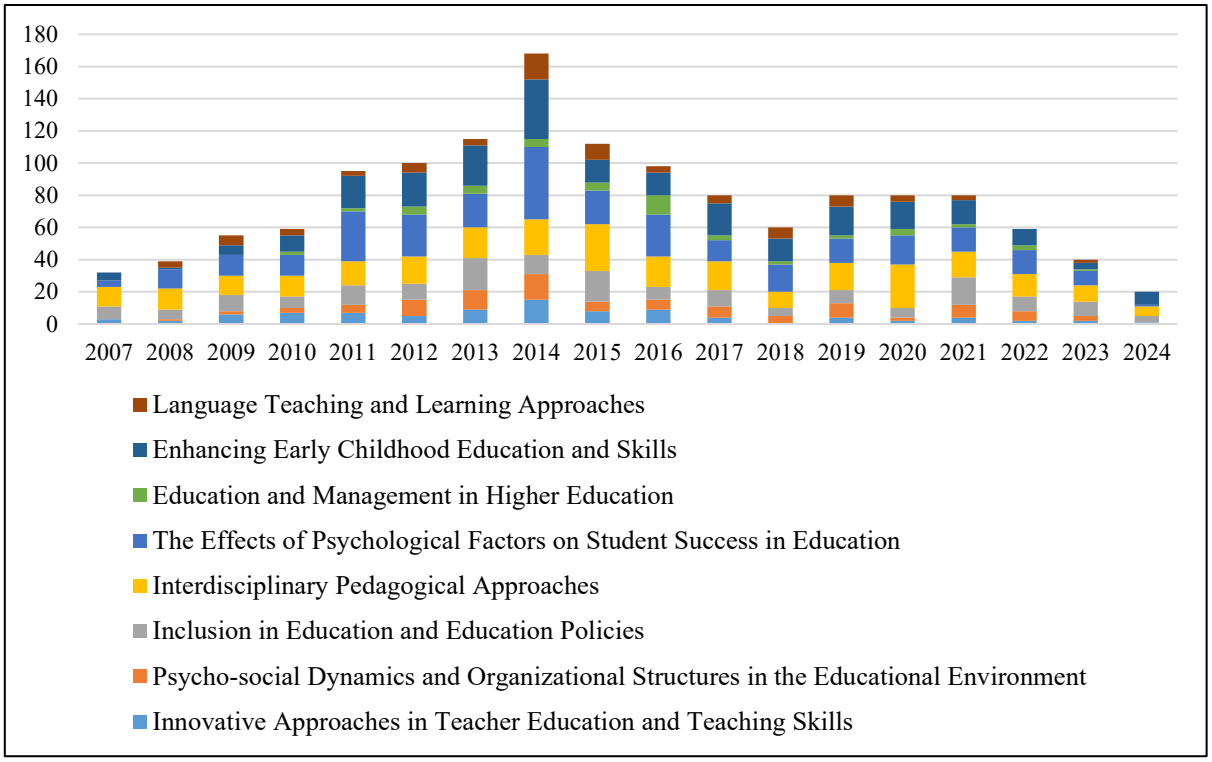
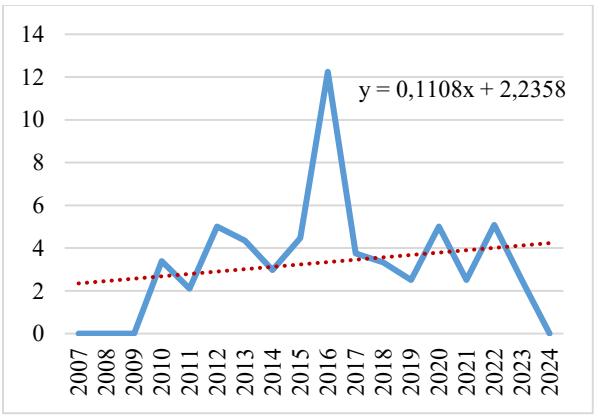
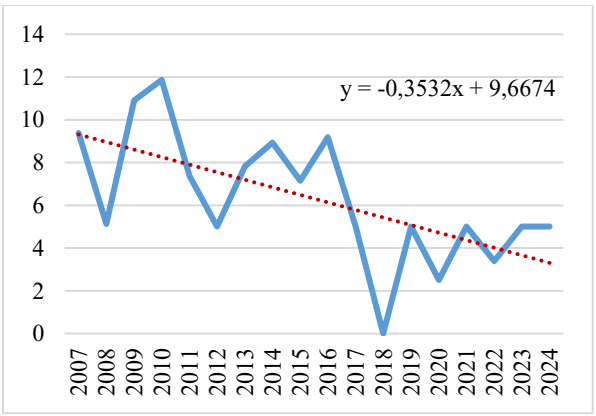


Figure 12. Distribution of the topics of the articles published in the Education and Science Journal

The volume of the topics in the Education and Science Journal has increased at some times and decreased at other times, especially between 2010 and 2015, when there was a significant increase in the number of articles in all topics (see Figure 12). When the trend lines of the topics are analyzed, it is seen that they remain constant horizontally. However, despite the temporal change in the Education and Science Journal, there is no change in terms of topic headings. Trend changes of the topics in the Education and Science Journal are given in Figure 13.



Education and Management in Higher Education



Innovative Approaches in Teacher Education and Teaching Skills

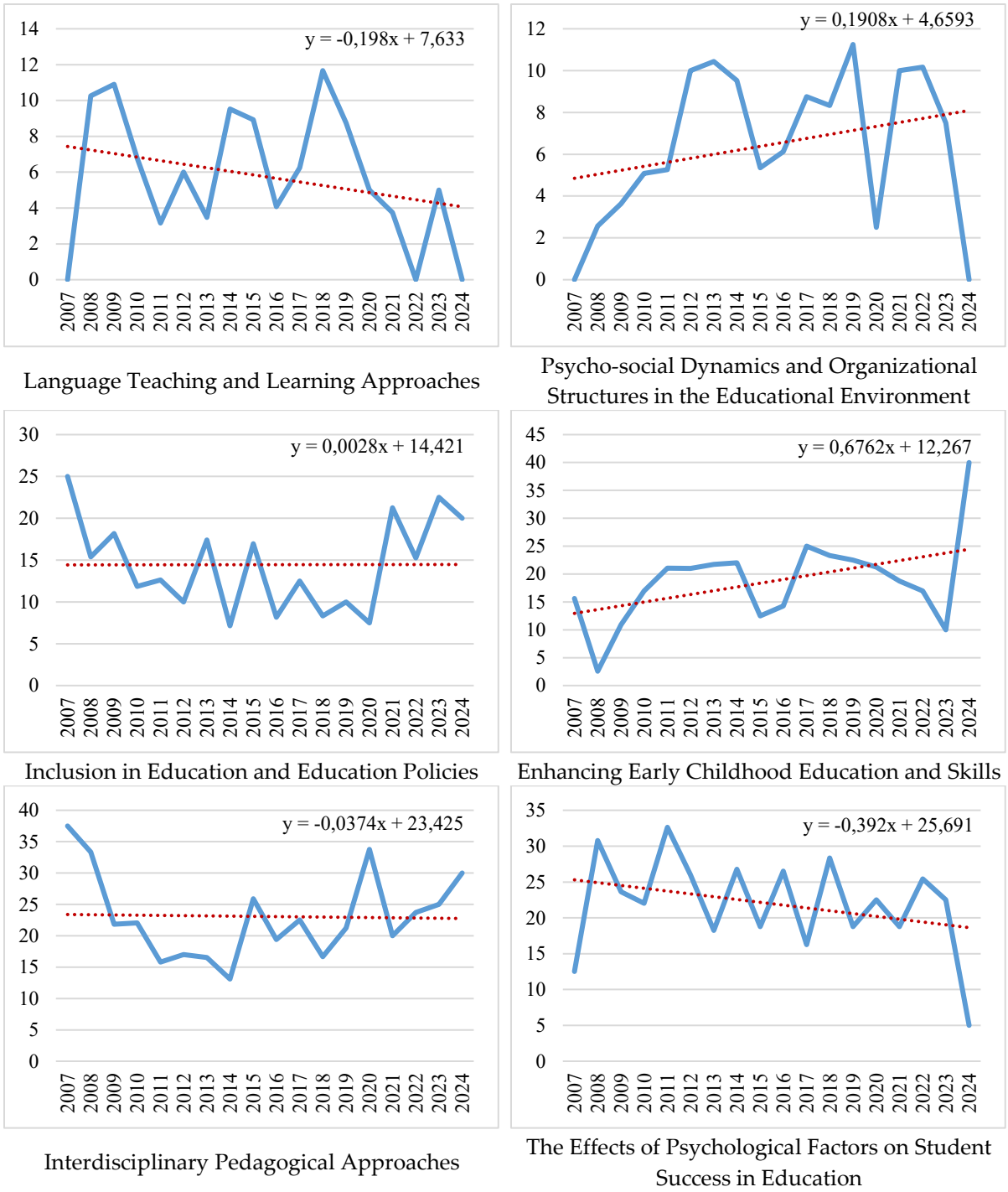


Figure 13. Trend changes in the topics covered by the articles published in the Education and Science Journal

Figure 13 depicts the trend changes in the topics over time. The trends of the topics in the Education and Science Journal either decrease, remain constant or increase over time. The trend slopes of the topics "Innovative Approaches in Teacher Education and Teaching Skills", "The Effects of Psychological Factors on Student Achievement in Education" and "Language Education and Learning Approaches" decrease; "Inclusion and Educational Policies in Education" and "Interdisciplinary Pedagogical Approaches" remain stable, while the topics "Psycho-social Dynamics and Organizational Structures in the Educational Environment", "Education and Management in Higher Education" and "Early Childhood Education and Increasing Skills" increase. However, when the slope graphs are analyzed, it cannot be said that the decreases or increases are very rapid.

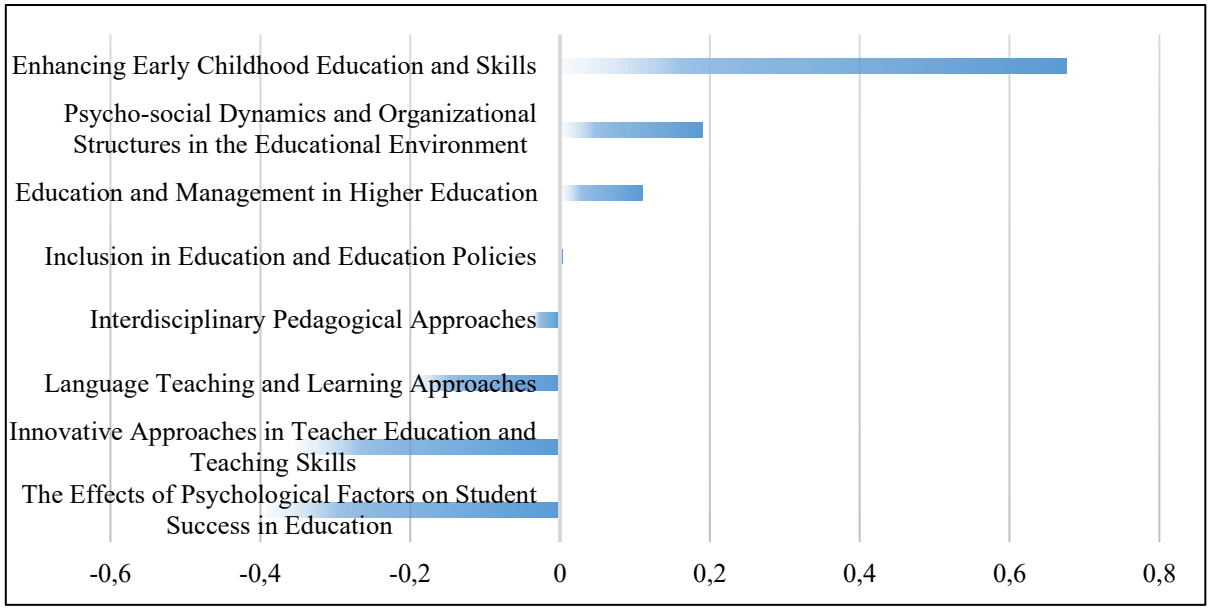


Figure 14. Momentum changes of the topics in the Education and Science Journal

Figure 14 highlights the changes in each topic in the Education and Science Journal compared to other topics. "Early Childhood Education and Increasing Skills", "Psycho-social Dynamics and Organizational Structures in the Educational Environment" and "Education and Management in Higher Education" have positive momentum values. Accordingly, it has been determined that these topics have gained relative importance compared to other topics. The acceleration value of "Inclusion in Education and Educational Policies" is close to zero. It cannot be said that there has been an increase or decrease in the importance of this topic. On the other hand, "Interdisciplinary Pedagogical Approaches", "Language Teaching and Learning Approaches", "Innovative Approaches in Teacher Education and Teaching Skills" and "The Effects of Psychological Factors in Education on Student Achievement" have negative values in terms of acceleration. It can be said that these topics have relatively lost their importance in the journal.

Discussion

In the previous section, we reported the results of comprehensive bibliometric analysis and topic modeling of articles published in the Education and Science Journal between 2007 and 2024 and answered the four research questions. In the following paragraphs, we summarize the results and highlight key insights. The sample size of n=1,273 articles are comparable to that of Altunışık (2023), who examined n=1,270 articles and Gülmez et al. (2021), who examined n=1,041 articles. It is also notably higher than Selçuk et al.'s (2014) analysis of n=492 articles and Doğan and Tok's (2018) study of n=181 articles published in the Education and Science Journal. As the most recent study, it integrates and confirms some earlier findings.

Compared to other journals in the field of educational research (Arici et al., 2019; Karagöz & Koç Ardıç, 2019; Karagöz & Şeref, 2019; Kutluca & Demirkol, 2016; Ünal, 2022) the Education and Science Journal publishes a higher-than-average number of articles, which allows it to be positioned as a significant contributor to educational research. However, due to a fluctuation in the annual article publication rates, the annual growth rate of the Education and Science Journal was found to be -2.73%. The analysis points to a decrease in the number of articles published by the Education and Science Journal in the recent period, which highlights an opportunity for improvement. The average age of the articles published in the journal is 8.77 years; it has been cited 7,390 times since being indexed by WoS. While the journal is mature and well-established at the national level, it is still emerging internationally. The fact that it is included in the Q4 (2024) /Q3 (2025) quartile in WoS denotes just that.

We identified Hacettepe University as the institution that publishes the most articles in the journal. Our results tally with those of Altunışık (2023), Doğan and Tok (2018), Gülmez et al. (2021), Selçuk et al. (2014), and Ünal (2022). These consistent results suggest that Hacettepe University is the leading institution in Turkey in the field of educational research. This may be attributed to the fact that Hacettepe University is a well-established institution. The results further highlighted two researchers from Hacettepe University who work in the field of measurement and evaluation in education. We have further noted the increased interest in articles in the field of measurement and evaluation. Moreover, in contrast to a study by Altunışık (2023), which identified Murat Özdemir from Hacettepe University as the most active researcher in the field, this finding was not confirmed by other studies including those by Doğan and Tok (2018), Gülmez et al. (2021) and Selçuk et al. (2014). However, the journal is primarily in demand from institutions in Turkey, as fewer papers are published by international institutions. This should be seen as part of the journal's ongoing improvement process, presenting a growth opportunity in pursuit of its vision to become a global leader.

Our analysis of the keywords also differed from that reported by Altunışık (2023). The most frequently used keywords in our sample were "teacher", "education", "validity", "reliability", "academic achievement", "gender", "self-efficacy", "social research", "higher education", and "teacher education". The infrequent use of terms such as "early childhood", "critical thinking", "leadership", "motivation", and "action research" suggests limited research on these topics. While there is some overlap in the keywords between Altunışık's (2023) paper and the present study, their frequencies were different. For example, the most frequently used keyword in Altunışık's (2023) study was "academic achievement" followed by "reliability, validity, teacher, gender, pre-service teachers, self-efficacy, higher education and structural equation modeling". In another analysis of the Education and Science Journal conducted by Ünal (2022), the most frequent keywords were "pre-service teacher, Covid-19, teacher, distance education, secondary school students and scale development". The variance may be attributed to the different methods used and changes in sample composition.

With respect to the countries represented by researchers in articles published in the Education and Science Journal; our results indicate that Turkey ranked first. This finding is consistent with that of Altunışık's (2023) and noted by Merigó et al. (2015). The result should not come as a surprise, as one would expect many researchers to prioritize journals published in their own language and country. However, international collaborations with Turkish co-authorship have been observed in our data. The country with the highest number of Turkish co-authors is the USA. The United States generally leads in educational science publications due to its extensive higher education system, research funding, and robust academic publishing infrastructure. The 2018 study titled "Scientific Publication Performance in Educational Sciences: G-20 Countries and Turkey," identifies the USA as the leading producer of educational science publications among G-20 nations (Selvitopu, 2018). Therefore, researchers tend to collaborate more with American colleagues.

The results of topic modeling similarly point to differences across studies. The most important topics identified in this study are teachers, education, academic achievement, validity and reliability, and gender. In a study by Selçuk et al. (2014), the key topics were the field of curriculum development and educational administration. In contrast, the most important topic identified by Altunışık (2023) was teacher opinions, followed by academic achievement, student perception, reading skills, teaching-learning conceptions and learning methods. Topic modeling conducted by Özcan and Akar (2024), though analyzing data from a different educational journal, identified educational sciences and administration as the main topics.

Topic modeling further identified eight topics. These include innovative approaches in teacher education and teaching skills, psycho-social dynamics and organizational structures in the educational environment, inclusion in education and education policies, interdisciplinary pedagogical approaches, the effects of psychological factors on student success in education, education and management in higher education, enhancing early childhood education and skills, and language teaching and learning approaches. These eight topics overlap with the focus and scope of the Journal of Education and Science, which aims to publish studies that contribute to educational activities at all levels, from pre-school education to higher education, and supports the professional development of teachers and academics (Eğitim ve Bilim, 2024). Researchers prefer the Journal of Education and Science due to its broad topical scope. The emerging topics align with the themes and key concepts discussed in the studies by Selçuk et al. (2014) and Altunışık (2023). In Turkey, teacher training and the continuous improvement of teaching skills are the responsibility of education departments affiliated with universities. We argue that it is important to discuss issues related to teacher education and the enhancement of teachers' professional skills in the Education and Science Journal.

The role of educational administration is critical not only for designing educational frameworks and drafting policies but also for maintaining the quality and effectiveness of the educational system. It is a key department in graduate programs at faculties of education, with the second-highest number of academics compared to other education fields (Council of Higher Education [CoHE], 2024). As a result, studies in educational administration offer valuable insights into how educational policies are developed and implemented, from the top levels of government to schools and classrooms. For educational administration to become a relevant discipline, it should strive to develop original approaches to a variety of managerial problems (Eacott, 2015). Studies in this field should take a pragmatic stance and address practical problems by considering the unique economic, political, cultural, and social characteristics and challenges of society, rather than repeating popular slogans (Takmak, 2019). The high ranking of educational administration research in the journal is likely a result of the number of publications in the field.

Inclusive education has emerged as an important topic and can be defined as a process that ensures all children receive education aligned with their developmental, sociocultural, and individual needs. The journal Education and Science (2024) also aims to promote inclusion in education. The inclusion of articles on this topic in the Education and Science Journal underscores the importance of inclusive education as an issue requiring urgent solutions. Increasing children's accessibility and fostering participation, while reducing potential discrimination are among the main goals of inclusive education according to (Ministry of National Education [MoNE], 2024). Furthermore, it has been repeatedly stressed that childhood experiences have a lasting impact on individuals' cognitive abilities, personalities, and social behaviors (Bredekamp, 2015; UNICEF, 2003). Some have noted that it is important to continuously renew assessment methods and tools to ensure the needs of child development are met (Tunçeli & Zembat, 2017).

Perhaps many of these issues could be addressed by employing interdisciplinary pedagogical approaches, identified as a topic in our analysis. Today, programs are no longer designed around a single subject area but rather combine multiple disciplines with the aim of providing hands-on experience. This is evident in recent curricula prepared by MoNE (2024), which emphasizes STEM and STEAM-based program and activity proposals (Özcan, 2021a; 2021b; 2023a; 2023b) and promotes a holistic education approach (Altan & Yıldırım, 2022). In particular, STEM applications have become a cornerstone of modern science education. Teaching STEM requires one to adopt an interdisciplinary approach, which enables its use across different ages and academic levels, from primary school to university (Özcan & Koca, 2019). The Journal of Education and Science notably lacks interdisciplinary studies. We argue that the inclusion of studies prepared with an interdisciplinary approach in the Education and Science Journal will increase the impact of the journal in literature.

Language education—an interdisciplinary field—cultivates both communicative competence and intercultural awareness. Communication-based approaches emphasize meaning-making over grammatical structure in language teaching (Richards, 2006). Furthermore, language education should not be confined to the narrow path of learning foreign languages. Turkey has developed several foreign language learning strategies (Soner, 2007), and the teaching of Turkish as a second language has gained increasing interest. Research into strategies for enhancing reading comprehension, a key issue identified in international assessments, should be addressed within the broader context of language learning. For this reason, the Journal of Education and Science should include not only studies on foreign language teaching, but also those that emphasize the importance of pedagogical content knowledge in language teaching, as well as studies that evaluate Turkish as a foreign language. Recognizing the importance of language education, the Journal of Education and Science features language teaching and learning approaches, helping to bridge a potential gap in interdisciplinary studies.

One may argue that higher education plays an important role in shaping a country's future by providing opportunities for personal growth and fostering its human potential. To develop a skilled and effective workforce, higher education policies must be designed with utmost care, proactively addressing challenges and providing opportunities for continuous improvement. To ensure success in global competition, higher education has gained increasing importance. This has led to rapid growth in the number of institutions, a trend seen in Turkey, where higher education has expanded significantly. Aydın et al. (2017) caution that an increase in quantity does not always lead to improved quality, and if the rise in quantity is not accompanied by quality, it can result in new problems. A quick search reveals that there are many academic journals in the field of educational sciences in Turkey. Although the exact number varies depending on the criteria and search period, one may estimate that there are approximately between 25 and 33 such journals (Danişman et al., 2016). Only one of these journals, Education and Science, is included in an internationally recognized index (WoS-SSCI). Other journals in the field are expected to be included in international indexes by following similar policies and publishing articles on a variety of topics.

Our findings also suggest that there are topics not covered by the Journal of Education and Science. These include lifelong learning, informal learning, artificial intelligence and the use of generative AI in education, pedagogical content knowledge, TPACK, educational comics, formative assessment and evaluation practices, educational policies developed at national and international level, and innovative education in the androgogy and gerontology process. We consider the absence of discourse on these topics as potential gaps that would be beneficial for the journal to address. Expanding the scope and inviting articles on these topics would contribute to the "focus journal" approach outlined in the journal's purpose statement.

Conclusion

In the previous sections we outlined the results of a bibliometric analysis and topic modeling (2007-2024) of n=1,372 papers published in the Education and Science Journal (Eğitim ve Bilim); a leading peer-reviewed journal published by the Turkish Education Association. It is the only journal in Turkey indexed in the SSCI category of the WoS, which publishes educational research in both Turkish and English languages. However, the journal is primarily engaged with by Turkish academics, rather than by researchers from other countries, although some international collaborations have been noted. This may explain the recent decline in the citation rate of the journal. We argue that the journal should take note and reposition itself as a global education journal rather than a national one.

As well, there are other reasons why it is the only journal from Turkey included in the WoS SSCI index. The wide range of topics covered in the journal's publication purpose and scope is seen as a reason for researchers to prefer it over others. The eight topics that emerged in the publication trend are an indication of this. Moreover, the journal's rigorous peer-review process, which ensures high-quality scholarship, undoubtedly contributes to its prestigious standing. The editorial board, composed of internationally recognized experts in their respective fields, further bolsters the journal's credibility and attracts submissions. Its commitment to publishing original research that advances knowledge, coupled with its accessibility to a broad audience, solidifies its position as a leading source of scholarly work and likely factors into its unique inclusion in the WoS SSCI index from Turkey.

To be indexed in the Social Sciences Citation Index (SSCI) by Clarivate Analytics, a journal must pass a rigorous evaluation (Clarivate Analytics, 2023). While the Journal of Education and Science has successfully completed this process, further improvements are needed to meet international standards and enhance its standing. To reduce bias and encourage submissions from new researchers and international scholars, the journal's publication policy should be updated with revised criteria, aims, and scope.

To mitigate the journal's limited growth rate, we suggest expanding its scope and publishing special issues on topics that are currently underrepresented, in order to increase the number of articles. We anticipate that this will effectively elevate the journal to a position where it can engage with educational issues on a global scale.

Limitations

The scope of the research, the database selected as the sample, the analyses performed, and the programs used in performing these analyses constitute the limitations of this study. The study was limited only to the articles indexed in WoS and published between the specified dates. Since the Education and Science Journal is still in the process of publication, new articles are being produced. With the inclusion of these articles in the study, the findings obtained may vary and the topics may differ. In addition, the use of the Bibliometrix package written in the R programming language in the bibliometric analysis part of the study and the application of the LDA technique in the topic modeling analysis are also limitations of this study in terms of analysis methods. Because the basic method in bibliometric analysis and topic modeling studies is to create models based on the frequency of words. It should be taken into consideration that the algorithms in these analysis programs do not always give complete and reliable results. The LDA technique for topic modeling analysis has recently been seen as important and up-to-date and has been popularly used in many studies. Its use in this study can be seen as a limitation. Therefore, different topic modeling techniques should be used, and the results should be compared.

References

Acar Güvendir, M. (2014). Öğrenci başarılarının belirlenmesi sınavında öğrenci ve okul özelliklerinin Türkçe başarısı ile ilişkisi. *Eğitim ve Bilim, 39*(172), 163-180. https://educationandscience.ted.org.tr/article/view/1256

Agada, J. (1987). Bibliometrics, history of the development of ideas. In A. Kent (Ed.), *Encyclopedia of library and information science* (Vol. 42, Suppl. 7, pp. 144-218). Marcel Dekker.

Aggarwal, C. C., & Zhai, C. (2013). *Mining text data.* Springer.

Aggarwal, R., & Ranganathan P. (2019). Study designs: Part 2- Descriptive studies. *Perspectives in Clinical Research, 10*(1), 34-36. https://doi.org/10.4103/picr.PICR_154_18

Ağaoğlu, E., Altınkurt, Y., Yılmaz, K., & Karaöse, T. (2012). Okul yöneticilerinin yeterliklerine ilişkin okul yöneticilerinin ve öğretmenlerin görüşleri (Kütahya ili). *Eğitim ve Bilim, 37*(164), 159-175. https://educationandscience.ted.org.tr/article/view/1035

Akman, Y., & Özdemir, M. (2018). Örgütsel çekicilik, örgütsel imaj ve örgütsel sadakat arasındaki ilişkilerin incelenmesi: Öğretmenler üzerine bir araştırma. *Eğitim ve Bilim, 44*(198), 1-16. https://doi.org/10.15390/EB.2018.7928

Alrayashi, A. M. (2023). *Analyzing higher education research trends with Latent Dirichlet Allocation: A text mining approach* (Thesis No. 812604) [Master's thesis, Atilim University]. Council of Higher Education National Thesis Center.

AlSumait, L., Barbará, D., & Domeniconi, C. (2008). On-line LDA: Adaptive topic models for mining text streams with applications to topic detection and tracking. In *Proceedings of 2008 Eighth IEEE International Conference on Data Mining* (pp. 3-12). IEEE. https://doi.org/10.1109/ICDM.2008.140

Altan, M. Z., & Yıldırım, T. (2022). Bütüncül eğitim, bütüncül gelişim, bütüncül insan. *Eğitimde Yeni Yaklaşımlar Dergisi, 5*(1), 68-77.

Altunışık, M. (2023). *Eğitim ve Bilim Dergisi'nin bibliyometrik ve içerik analizi* (Thesis No. 777707) [Master's thesis, Necmettin Erbakan University]. Council of Higher Education National Thesis Center.

Aria, M., & Cuccurullo, C. (2017). Bibliometrix: An R-tool for comprehensive science mapping analysis. *Journal of Informetrics, 11*(4), 959-975. https://doi.org/10.1016/j.joi.2017.08.007

Arici, F., Yıldırım, P., Çalıklar, Ş., & Yılmaz, R. M. (2019). Research trends in the use of augmented reality in science education: Content and bibliometric mapping analysis. *Computers & Education, 142*, 103647. https://doi.org/10.1016/j.compedu.2019.103647

Arslan, G. (2016). Lise öğrencilerinde reddedilmişlik duygusu, akademik başarı, akademik yeterlik ve eğitimsel amaçlar arasındaki ilişki. *Eğitim ve Bilim, 41*(183), 293-304. https://doi.org/10.15390/EB.2016.5562

Aslan, G. (2021). Eğitime erişimin belirleyicileri: Kırsal kesimdeki kız çocukların ortaöğretime gönderilmeme nedenleri. *Eğitim ve Bilim, 46*(207), 169-201. https://doi.org/10.15390/EB.2021.9045

Atik, S., & Özer, N. (2020). Lise öğrencilerinin öğretmene güvenleri ile akademik başarıları arasındaki ilişkide okula karşı tutum okula yabancılaşma ve okul tükenmişliğinin doğrudan ve dolaylı rolü. *Eğitim ve Bilim, 45*(202), 441-458. https://doi.org/10.15390/EB.2020.8500

Aydın, M. S., Kaya, H., & Kete, H. (2017). Yükseköğretim hizmetlerinin özellikleri ve Türkiye'de yükseköğretim hizmetlerinin görünümü. *Journal of Awareness, 2*(2), 1-24.

Baran, B., & Ata, F. (2013). Üniversite öğrencilerinin web 2.0 teknolojileri kullanma durumları, beceri düzeyleri ve eğitsel olarak faydalanma durumları. *Eğitim ve Bilim, 38*(169), 192-208. https://educationandscience.ted.org.tr/article/view/1166

Barde, B. V., & Bainwad, A. M. (2017, June) An overview of topic modeling methods and tools. In *International Conference on Intelligent Computing and Control Systems (ICICCS)* (pp. 745-750). IEEE. https://doi.org/10.1109/ICCONS.2017.8250563

Bican, G., & Demir, N. (2018). Danimarka'da Türkçe-Danca konuşan öğrencilerin dil baskınlığının iki dillilik temelinde incelenmesi. *Eğitim ve Bilim, 43*(194), 131-155. https://doi.org/10.15390/EB.2018.7418

Blei, D. M., Edu, B. B., Ng, A. Y., Edu, A. S., Jordan, M. I., & Edu, J. B. (2003). Latent dirichlet allocation. *Journal of Machine Learning Research, 3*, 993-1022.

Bredekamp, S. (2015). *Erken çocukluk eğitiminde etkili uygulamalar* (2nd ed., H. Z. İnan & T. İnan, Trans.). Nobel Yayıncılık.

Buyruk, H., & Akbaş, A. (2021). Öğretmenlerin mesleki profesyonelliği ile özerkliği arasındaki ilişkiye dair bir çözümleme. *Eğitim ve Bilim, 46*(208), 431-451. https://doi.org/10.15390/EB.2021.9996

Büyükşahin Çevik, G. (2017). Okul yöneticilerinin empatik eğilimlerini yordamada öfke, yaşam doyumu ve stresle başa çıkmanın rolü. *Eğitim ve Bilim, 42*(191), 283-305. https://doi.org/10.15390/EB.2017.7039

Bystrov, V., Naboka-Krell, V., Staszewska-Bystrova, A., & Winker, P. (2023). *Analysing the impact of removing infrequent words on topic quality in LDA models*. arXiv. https://arxiv.org/abs/2311.14505

Cao, J., Xia, T., Li, J., Zhang, Y., & Tang, S. (2009). A density-based method for adaptive LDA model selection. *Neurocomputing, 72*(7-9), 1775-1781. https://doi.org/10.1016/j.neucom.2008.06.01

Chen, X., Zou, D., & Xie, H. (2020). Fifty years of British Journal of educational technology: A topic modeling based bibliometric perspective. *British Journal of Educational Technology, 51*(3), 692-708. https://doi.org/10.1111/bjet.12907

Cho, Y. J., Fu, P. W., & Wu, C. C. (2017). Popular research topics in marketing journals, 1995-2014. *Journal of Interactive Marketing, 40*(1), 52-72. https://doi.org/10.1016/j.intmar.2017.06.003

Clarivate Analytics. (2023). *Web of Science journal evaluation process and selection criteria.* https://clarivate.com/webofsciencegroup/journal-evaluation-process-and-selection-criteria/

Council of Higher Education. (2024). *Birim bazında öğretim elemanı sayıları raporu.* https://istatistik.yok.gov.tr/

Danişman, Ş., Yalçın, M., Çiftçi, Ş. K., Tosuntaş, Ş. B., Sölpük, N., Ay, Y., Karadağ, E., & Yücel, C. (2016). Scientific publication map of Journals published in the field of educational sciences in Turkey: An analysis impact factors of journals. *Educational Administration Theory and Practice Journal, 22*(4), 483-506. https://doi.org/10.14527/kuey.2016.019

Dao, L. T., Tran, T., Van Le, H., Nguyen, G. N., & Trinh, T. P. T. (2023). A bibliometric analysis of research on Education 4.0 during the 2017-2021 period. *Education and Information Technologies, 28*(3), 2437-2453. https://doi.org/10.1007/s10639-022-11211-4

Demir Başaran, S. (2020). Suriyeli mülteci öğrencilerin öğretmeni olmak: Öğretmenlerin okul deneyimleri. *Eğitim ve Bilim, 46*(206), 331-354. https://doi.org/10.15390/EB.2020.9182

Demirtaş, H., Cömert, M., & Özer, N. (2011). Öğretmen adaylarının özyeterlik inançları ve öğretmenlik mesleğine ilişkin tutumları. *Eğitim ve Bilim, 36*(159), 96-111. https://educationandscience.ted.org.tr/article/view/910

Demšar, J., Curk, T., Erjavec, A., Gorup, Č., Hočevar, T., Milutinovič, M., Možina, M., Polajnar, M., Toplak, M., Starič's, A., Štajdohar, M., Umek, L., Žagar, L., Žbontar, J., Žitnik, M, & Zupan, B. (2013). Orange: data mining toolbox in Python. *The Journal of Machine Learning Research, 14*(1), 2349-2353.

Dimici, K., Seggie, F., Hacifazlioglu, O., & Caner, A. (2016). Yüksek öğretimde bölüm başkanlarının profesyonelleşmelerindeki zorluklar: Türkiye'de bir nitel çalışma. *Eğitim ve Bilim, 41*(184), 131-146. https://doi.org/10.15390/EB.2016.6149

Doğan, H., & Tok, T. N. (2018). Türkiye'de eğitim bilimleri alanında yayınlanan makalelerin incelenmesi: Eğitim ve Bilim Dergisi örneği. *Current Research in Education, 4*(2), 94-109. https://doi.org/10.38015/sbyy.1180249

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

Eacott, S. (2015). The principalship, autonomy, and after. *Journal of Educational Administration and History, 47*(4), 414-431. https://doi.org/10.1080/00220620.2015.996866

Eğitim ve Bilim. (2024). *Dergi hakkında.* https://educationandscience.ted.org.tr/about

Erkuş, A. (2021). *Davranış bilimleri için bilimsel araştırma süreci* (7th ed.). Seçkin.

Erten, P. (2022). Meslek liselerinin 21. yy. öğrenen ve öğreten becerileri kapsamında değerlendirilmesi. *Eğitim ve Bilim, 47*(209), 261-291. https://doi.org/10.15390/EB.2022.10702

Ferguson, R., Barzilai, S., Ben-Zvi, D., Chinn, C., Herodotou, C., Hod, Y., Kali, Y., Kukulska-Hulme, A., Kupermintz, H., & McAndrews, P. (2017). *Innovating pedagogy 2017.* OpenUniv.

Genç, Z. (2014). İngilizce öğretiminde sözlü iletişim hatalarının düzeltilmesi: Farklı dil düzeylerindeki Türk öğrencilerin tercihleri. *Eğitim ve Bilim, 39*(174), 259-271. https://doi.org/10.15390/EB.2014.1438

Google Scholar. (2024). *Google Scholar metrcis*. https://scholar.google.com/citations?hl=tr&view_op=search_venues&vq=e%C4%9Fitim+ve+bilim&btnG=

Gülmez, D., Özteke, İ., & Gümüş, S. (2021). Overview of educational research from Turkey published in international journals: A bibliometric analysis. *Eğitim ve Bilim, 46*(206), 213-239. https://doi.org/10.15390/eb.2020.9317

Gültekin Toroslu, A. (2013). Halk eğitimi merkezlerinde uygulanmakta olan modüler yetişkin eğitimi programlarının yeterliliklerinin araştırılması. *Eğitim ve Bilim, 38*(168), 3-16. https://educationandscience.ted.org.tr/article/view/1123

Gündoğdu, K. (2011). Türkiye'de öğretmen adaylarının insan hakları eğitimine yönelik tutumları. *Eğitim ve Bilim, 36*(162), 182-195. https://educationandscience.ted.org.tr/article/view/987

Gür, Ç., Koçak, N., & Sağlar, M. (2017). Çok boyutlu bakış açılarıyla düşünme eğitimi programının 5 ve 6 yaş grubu çocukların okul olgunluğu üzerindeki etkisinin incelenmesi. *Eğitim ve Bilim, 42*(189), 153-165. https://doi.org/10.15390/EB.2017.6760

Hallinger, P., & Kovačević, J. (2019). A bibliometric review of research on educational administration: Science mapping the literature, 1960 to 2018. *Review of Educational Research, 89*(3), 335-369. https://doi.org/10.3102/0034654319830380

Hotaman, D., & Yüksel Şahin, F. (2010). Öğretim elemanı coşkusunun üniversite öğrencilerinin başarı güdüsü düzeyleri üzerindeki etkisi. *Eğitim ve Bilim, 35*(155), 89-103. https://educationandscience.ted.org.tr/article/view/851

Huang, C., Yang, C., Wang, S., Wu, W., Su, J., & Liang, C. (2020). Evolution of topics in education research: A systematic review using bibliometric analysis. *Educational Review, 72*(3), 281-297. https://doi.org/10.1080/00131911.2019.1566212

Indriyani, V., Atmazaki, A., & Ramadhan, S. (2023). Dil öğreniminde çoklu okuryazarlık temelli bütünleştirici öğrenme (MULGRANING) modelinin geliştirilmesi. *Eğitim ve Bilim, 48*(215), 261-275. https://doi.org/10.15390/EB.2023.11413

İlhan, G., & Oruç, Ş. (2019). Sosyal bilgiler dersinde çizgi roman kullanımı: Teksas tarihi. *Eğitim ve Bilim, 44*(198), 327-341. https://doi.org/10.15390/EB.2019.7830

Johri, A., Joshi, P., Kumar, S., & Joshi, G. (2024). Metaverse for sustainable development in a bibliometric analysis and systematic literature review. *Journal of Cleaner Production, 435*, 140610. https://doi.org/10.1016/j.jclepro.2024.140610

Karababa, C., & Karagül, S. (2013). Yabancı dil olarak Türkçe öğrenenlere yönelik gereksinim çözümlemesi. *Eğitim ve Bilim, 38*(170), 361-371. https://educationandscience.ted.org.tr/article/view/1208

Karagöz, B., & Koç Ardıç, İ. (2019). Ana Dili Eğitimi Dergisinde yayımlanan makalelerin bibliyometrik analizi. *Ana Dili Eğitimi Dergisi, 7*(2), 419-435. https://doi.org/10.16916/aded.482628

Karagöz, B., & Şeref, İ. (2020) Yazma becerisiyle ilgili makaleler üzerine bir inceleme: Web of Science veri tabanında eğilimler. *Ana Dili Eğitimi Dergisi, 8*(1), 67-86. https://doi.org/10.16916/aded.619090

Kavacık, L., Yanpar Yelken, T., & Sürmeli, H. (2015). İlköğretim fen ve teknoloji dersinde inovasyon (yenilikçi) proje uygulamaları ve öğrenciler üzerindeki etkileri. *Eğitim ve Bilim, 40*(180), 247-263. https://doi.org/10.15390/EB.2015.2613

Kılınç, A. (2014). Öğretmen profesyonelizminin bir yordayıcısı olarak okul kültürü. *Eğitim ve Bilim, 39*(174), 105-118. https://doi.org/10.15390/EB.2014.2756

Köroğlu, S. A. (2015). Literatür taraması üzerine notlar ve bir tarama tekniği. *GİDB Dergi,* (01), 61-69.

Kutluca, T., & Demirkol, M. (2016). Dicle Üniversitesi Ziya Gökalp Eğitim Fakültesi dergisinin bibliyometrik analizi. *Dicle Üniversitesi Ziya Gökalp Eğitim Fakültesi Dergisi*, (28), 108-118. https://doi.org/10.14582/duzgef.674

Lewis, C. M., & Grossetti, F. (2022). A statistical approach for optimal topic model identification. *Journal of Machine Learning Research, 23*(58), 1-20.

Li, K., Rollins, J., & Yan, E. (2018). Web of science use in published research and review papers 1997-2017: A selective, dynamic, cross-domain, content-based analysis. *Scientometrics, 115*, 1-20. https://doi.org/10.1007/s11192-017-2622-5

Merigó, J. M., Mas-Tur, A., Roig-Tierno, N., & Ribeiro-Soriano, D. (2015). A bibliometric overview of the Journal of Business Research between 1973 and 2014. *Journal of Business Research, 68*(12), 2645-2653. https://doi.org/10.1016/j.jbusres.2015.04.006

Mimno, D., Wallach, H., Talley, E., Leenders, M., & McCallum, A. (2011, July). Optimizing semantic coherence in topic models. In R. Barzilay & M. Johnson (Eds.), *Proceedings of the 2011 Conference on Empirical Methods in Natural Language Processing* (pp. 262-272). Association for Computational Linguistics. https://aclanthology.org/D11-1024.pdf

Ministry of National Education. (2024). *Kapsayıcı eğitim*. https://egitimdebirlikteyiz.meb.gov.tr/ContentPage/KapsayiciEgitim

Mirici, İ., İlter, B., Saka, Ö., & Glover, P. (2009). Eğitim amaçlı değişim programları ve Erasmus yoğun dil kursları: Türkçe kursları için bir örnek olay. *Eğitim ve Bilim, 34*(152), 148-159. https://educationandscience.ted.org.tr/article/view/814

Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & Prisma Group. (2009). Reprint-preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *Physical Therapy, 89*(9), 873-880. https://doi.org/10.1093/ptj/89.9.873

Özberk, E. B. Ü., Fındık, L. Y., & Özberk, E. H. (2018). Investigation of the variables affecting the math achievement of resilient students at school and student level. *Eğitim ve Bilim, 43*(194), 111-129. https://doi.org/10.15390/eb.2018.7153

Özcan, H. (2021a). *STEM eğitimi uygulamaları-1*. Pusula 20 Teknoloji ve Yayıncılık A.Ş.

Özcan, H. (2021b). *STEM eğitimi uygulamaları-2*. Pusula 20 Teknoloji ve Yayıncılık A.Ş.

Özcan, H. (2023a). *STEM eğitimi uygulamaları-3*. Vizetek Yayıncılık.

Özcan, H. (2023b). *STEM eğitimi uygulamaları-4*. Vizetek Yayıncılık.

Özcan, H., & Koca, E. (2019). STEM'e yönelik tutum ölçeğinin Türkçeye uyarlanması: Geçerlik ve güvenirlik çalışması. *Hacettepe Üniversitesi Eğitim Fakültesi Dergisi, 34*(2), 387-401. https://doi.org/10.16986/HUJE.2018045061

Özcan, H. Z., & Akar, C. (2024). Ulusal Eğitim Dergisi'nde yayınlanan makalelerin bibliyometrik analizi. *Ulusal Eğitim Dergisi, 4*(1), 456-467. https://doi.org/10.5281/zenodo.10605971

Özgün, S., Özkul, B., Oral, E., & Şemin, İ. (2020). Yoga eğitiminin erken çocukluk dönemindeki çocukların bilişsel işlevlerine etkisi. *Eğitim ve Bilim, 46*(206), 303-316. https://doi.org/10.15390/EB.2020.9088

Özkan, E. (2015). *Topic modelling by using Latent Dirichlet Allocation* (Thesis No. 518310) [Master's thesis, University of Nottingham]. Council of Higher Education National Thesis Center.

Özyurt, Ö., & Ayaz, A. (2022). Twenty-five years of education and information technologies: Insights from a topic modeling based bibliometric analysis. *Education and Information Technologies, 27*(8), 11025-11054. https://doi.org/10.1007/s10639-022-11071-y

Polat, Ö. (2021). 60-72 aylık çocukların yaratıcılık becerileri üzerinde zihin haritalama çalışmalarının etkisi. *Eğitim ve Bilim, 46*(207), 21-42. https://doi.org/10.15390/EB.2021.9582

Richards, J. C. (2006). *Communicative language teaching today.* Cambridge University Press.

Roemer, R. C., & Borchardt, R. (2015). *Meaningful metrics: A 21st century librarian's guide to bibliometrics, altimetrics, and research impact.* Amer Library Assn.

Rojas-Sánchez, M. A., Palos-Sánchez, P. R., & Folgado-Fernández, J. A. (2023). Systematic literature review and bibliometric analysis on virtual reality and education. *Education and Information Technologies, 28*(1), 155-192. https://doi.org/10.1007/s10639-022-11167-5

Rouse, M. (2018). *Text mining (Text analytics).* https://www.techtarget.com/searchbusinessanalytics/definition/text-mining

Selçuk, Z., Palancı, M., Kandemir, M., & Dündar, H. (2014). Eğitim ve Bilim dergisinde yayınlanan araştırmaların eğilimleri: İçerik analizi. *Eğitim ve Bilim, 39*(173), 430-453. https://educationandscience.ted.org.tr/article/view/1290

Selekler-Gökşen, N., Yıldırım Öktem, Ö., & İnelmen, K. (2016). Devlet ve vakıf üniversitelerinde örgütsel adalet algısının lider-üye etkileşimi üzerindeki etkisi. *Eğitim ve Bilim, 41*(184), 383-398. https://doi.org/10.15390/EB.2016.6139

Selvitopu, A. (2018). Eğitim bilimleri alanı bilimsel yayın performansı: G-20 ülkeleri ve Türkiye. *Journal of Turkish Studies, 13*(15), 383-399. https://doi.org/10.7827/turkishstudies.13379

Shadrova, A. (2021). Topic models do not model topics: Epistemological remarks and steps towards best practices. *Journal of Data Mining & Digital Humanities, 6*(2), 35-51. https://doi.org/10.46298/jdmdh.7595

Soner, O. (2007). Türkiye'de yabanci dil eğitiminin dünü bugünü. *Öneri Dergisi, 7*(28), 397-404. https://doi.org/10.14783/maruoneri.684553

Susar Kırmızı, F. (2015). Yaratıcı drama ve yaratıcı yazma uygulamalarının yaratıcı yazma başarısına etkisi. *Eğitim ve Bilim, 40*(181), 93-115. https://doi.org/10.15390/EB.2015.4552

Szomszor, M. (2024). *Introducing the Journal Citation Indicator: A new, field-normalized measurement of journal citation impact*. Clarivate. https://clarivate.com/blog/introducing-the-journal-citation-indicator-a-new-field-normalized-measurement-of-journal-citation-impact/

Şirin, H. (2009). Sivil toplum örgütlerinin eğitime ilişkin karar alma süreçlerine katılımları üzerine bir araştırma. *Eğitim ve Bilim, 34*(153), 169-182. https://educationandscience.ted.org.tr/article/view/831

Takmak, H. (2019). The originality and autonomy problem of educational administration. *Ege Eğitim Dergisi, 20*(1), 113-126. https://doi.org/10.12984/egeefd.452705

Tang, Y. K., Mao, X. L., Huang, H., Shi, X., & Wen, G. (2018). Conceptualization topic modeling. *Multimedia Tools and Applications*, 77, 3455-3471. https://doi.org/10.48550/arXiv.1704.02090

Temli, Y., Şen, D., & Akar, H. (2013). İlköğretim okulu öğretmen adaylarının ahlak ve ahlak eğitimi algıları ve tanımları. *Eğitim ve Bilim, 38*(168), 198-214. https://educationandscience.ted.org.tr/article/view/1137

Tozduman Yaralı, K., & Güngör Aytar, F. (2020). Öyküleştirme yöntemiyle oluşturulan eğitim programının okul öncesi çocukların eleştirel düşünme becerileri üzerindeki etkililiği. *Eğitim ve Bilim, 46*(205), 137-159. https://doi.org/10.15390/EB.2020.8698

Tunçeli, H. İ., & Zembat, R. (2017). Erken çocukluk döneminde gelişimin değerlendirilmesi ve önemi. *Eğitim Kuram ve Uygulama Araştırmaları Dergisi, 3*(3), 01-12.

Tyagi, N. (2021). *What is text mining? Process, methods and applications.* https://www.analyticssteps.com/blogs/what-text-mining-process-methods-and-applications

UNICEF. (2003). *The state of the world's children 2003.* https://digitallibrary.un.org/record/483509?v=pdf

Ursavaş, Ö., Şahin, S., & Mcilroy, D. (2014). Türkiye'deki öğretmen adaylarının BİT kullanımına yönelik davranışsal niyetlerinin belirlenmesinde branşlarının rolü. *Eğitim ve Bilim, 39*(175), 136-153. https://doi.org/10.15390/EB.2014.3357

Uslu, B. (2016). Modern üniversite ortamında öğretim üyelerinin akademik entelektüel liderliği. *Eğitim ve Bilim, 41*(184), 193-211. https://doi.org/10.15390/EB.2016.6140

Uysal, A. K., & Gunal, S. (2014). The impact of preprocessing on text classification. *Information Processing and Management, 50*(1), 104-112. https://doi.org/10.1016/j.ipm.2013.08.006

Uzun, A., Gökkurt Özdemir, B., Çetinöz, E., & Şahan, G. (2022). Hafif düzeyde zihinsel yetersizliği olan bir öğrenciye sesli düşünme yöntemi ile çıkarma işleminin öğretimi. *Eğitim ve Bilim, 47*(210), 321-356. https://doi.org/10.15390/EB.2022.10491

Ünal, M., & Aral, N. (2014). Fen eğitiminde problem çözme ölçeğinin geliştirilmesi: Geçerlik ve güvenirlik çalışmaları. *Eğitim ve Bilim, 39*(176), 267-278. https://doi.org/10.15390/EB.2014.3585

Ünal, O. (2022). Türkiye'de eğitim fakültesi dergilerinde yayınlanan makalelerin bibliyometrik analizi. *International Journal of New Approaches in Social Studies, 6*(2), 256-281. https://doi.org/110.38015/sbyy.1180249

Üstünlüoğlu, E. (2016). Algılar ve gerçekler: Yüksek öğretimde öğretim kalitesi. *Eğitim ve Bilim, 41*(184), 235-250. https://doi.org/10.15390/EB.2016.6095

Vayansky, I., & Kumar, S. A. (2020). A review of topic modeling methods. *Information Systems, 94*, 101582. https://doi.org/10.1016/j.is.2020.101582

Wallen, N. E., & Fraenkel, J. R. (2013). *Educational research: A guide to the process.* Taylor & Francis.

Yin, B., & Yuan, C. (2022). Detecting latent topics and trends in blended learning using LDA topic modeling. *Education and Information Technologies, 27*(9), 12689-12712. https://doi.org/10.1007/s10639-022-11118-0

Yun, E. (2020). Review of trends in physics education research using topic modeling. *Eğitim ve Bilim, 19*(3), 388-400. https://doi.org/10.33225/jbse/20.19.388

Zayimoğlu Öztürk, F. (2021). Eğitim sisteminde mülteciler: Türk ve mülteci öğrenci görüşleri. *Eğitim ve Bilim, 46*(208), 321-348. https://doi.org/10.15390/EB.2021.10439

Zhou, Y., & Wu, J. (2016). Oyun planı: Küresel Güney'in dünya standartlarındaki üniversitelerinin gelişimindeki dört karşıtlık. *Eğitim ve Bilim, 41*(184), 75-89. https://doi.org/10.15390/EB.2016.6152

Zupic, I., & Cater, T. (2015). Bibliometric methods in management and organization. *Organizational Research Methods, 18*(3), 429-472. https://doi.org/10.1177/1094428114562629