

international
electronic journal of
**elementary
education**



INTERNATIONAL ELECTRONIC JOURNAL OF ELEMENTARY EDUCATION

Editor in Chief

Kamil ÖZERK
University of Oslo, Norway

Editors

Gökhan ÖZSOY
Ordu University, Turkey

Annemie DESOETE
*Ghent University,
Arteveldehogeschool, Sig, Belgium*

Karen M. ZABRUCKY
Georgia State University, United States

Kathy HALL
University College Cork, Ireland

Turan TEMUR
Anadolu University, Turkey

Murat Doğan ŞAHİN
Anadolu University, Turkey

Hayriye Gül KURUYER
Ordu University, Turkey

Editorial Assistants

Abdullah KALDIRIM
Dumlupınar University, Turkey

Graphic Design

Vedat ŞEKER
*Kahramanmaraş Sutcu Imam University,
Turkey*

International Advisory Board

Bracha KRAMARSKI, Bar Ilan University, Israel

Collin Robert BOYLAN, Charles Sturt University, Australia

David Warwick WHITEHEAD, The University of Waikato, New Zealand

Dawn HAMLIN, SUNNY Oneonta, United States

Wendy HARRIOTT, Monmouth University, United States

Isabel KILLORAN, York University, Canada

Janelle Patricia YOUNG, Australian Catholic University, Australia

Jeanne ROLIN-IANZITI, The University of Queensland, Australia

Janet ALLEN, United States

Kouider MOKHTARI, Iowa State University, United States

Lloyd H. BARROW, University of Missouri, United States

Lori G. WILFONG, Kent State University, United States

Maria Lourdes DIONISIO, University of Minho, Portugal

Maribel GARATE, Gallaudet University, United States

Peter JOONG, Nipissing University, Canada

Ruth REYNOLDS, University of Newcastle, Australia

Therese Marie CUMMING, University of New South Wales, Australia

ISSN: 1307-9298

www.iejee.com
iejee@iejee.com



Education
&
Publishing



**All responsibility for statements made or opinions expressed in articles
lies with the author.**



This page is intentionally left blank.
www.iejee.com

Table of Contents

Effectiveness of Smart ABA in Teaching Reading Skills to Children with Autism Rudy Sutadi, Arneliza, Raihana, Asmadi Alsa, Kuncono Teguh Yunanto, Alam Zeb Khattak	697-710
The Relationship Between Parenting Styles and Perceived Stress Among Primary School Students: The Mediating Role of Creative and Critical Thinking Dispositions Sevil Momeni Shabani, Gülten Feryal Gündüz, Levent Vural	711-727
Integrating Water Sustainability into Primary Science Education: Teachers' Pedagogical Ration- ales, Opportunities, and Barriers Puji Winarti, Sri Haryani, Sri Wardani, Langlang Handayani	729-744
A Case Study of Kindergarten Teachers' Science Practices: Reflections of Knowledge and Attitude Nefise Özok Bulut, Meral Ören	745-759
Using Social Media for Participant Recruitment in Early Childhood Education and Care Research: A Reflective Analysis Elysia Phillips, Margaret Sims, Nicola Schutte, Marg Rogers	761-768
Parent Engagement in Early Childhood Education: A Multivariate Analysis in the Turkish Context Ümit Ünsal Kaya, Ümmühan Akpınar Afşin	769-784
The Development and Validation of the Metacognitive Awareness Scale for Middle School Students Ayşegül İşlekeller Bozca, Ayşe Gönültaş Yıldırım	785-799
Reframing the Reasoning Gap: Cognitive-Domain Achievement Profiles and Their Predictors Among Eighth-Grade Students in Türkiye Across TIMSS 2019 and 2023 Başak Barak, Betül Barut	801-813
Enhancing Elementary Students' Literacy Engagement through a Web-Based Digital Literature Platform: A Mixed-Methods Study and Instructional Framework Nurjanah, Yurdayanti, Yorenza Meifinda, Ilpandari, Winda Purnama Sari	815-837

Effectiveness of Smart ABA in Teaching Reading Skills to Children with Autism

Rudy Sutadi^a, Arneliza^{b,*}, Raihana^c, Asmadi Alsa^d, Kuncono Teguh Yunanto^e, Alam Zeb Khattak^f

Received : 21 October 2025
 Revised : 21 June 2026
 Accepted : 16 July 2026
 DOI : 10.26822/iejee.2026.460

^aRudy Sutadi, Faculty of Psychology, Persada Indonesia University- UPI Y.A.I., Jakarta, Indonesia.
 E-mail: rudysutadi123@gmail.com
 ORCID: <https://orcid.org/0000-0002-4437-8852>

^{b,*} **Corresponding Author:** Arneliza, Faculty of Psychology, Persada Indonesia University- UPI Y.A.I., Jakarta, Indonesia.
 E-mail: arneliza.2366290016@upi-yai.ac.id
 ORCID: <https://orcid.org/0009-0000-2205-2838>

^cRaihana, Faculty of Psychology, Persada Indonesia University- UPI Y.A.I., Jakarta, Indonesia.
 E-mail: raihana.2166290012@upi-yai.ac.id
 ORCID: <https://orcid.org/0000-0002-7078-5793>

^dAsmadi Alsa, Faculty of Psychology, Persada Indonesia University- UPI Y.A.I., Jakarta, Indonesia.
 E-mail: asmals@ugm.ac.id
 ORCID: <https://orcid.org/0000-0002-4222-4825>

^eKuncono Teguh Yunanto, Faculty of Psychology, Persada Indonesia University- UPI Y.A.I., Jakarta, Indonesia.
 E-mail: kunconoyunanto@gmail.com
 ORCID: <https://orcid.org/0000-0003-0168-0995>

^fAlam Zeb Khattak, School of Business Administration, Psychology Program IQRA National University Peshawar, Pakistan.
 E-mail: alamzeb@iqra.edu.pk
 ORCID: <https://orcid.org/0000-0001-8601-7089>

Abstract

Autism Spectrum Disorder is a neurodevelopmental disorder that appears before the age of three, marked by persistent deficits in communication, social interaction, and learning abilities. Reading is an essential skill for children with Autism Spectrum Disorder (ASD), supporting their academic and daily functioning. This study examined the effectiveness of Smart Applied Behavior Analysis (Smart ABA) in teaching foundational reading skills to children with ASD. Five nonverbal children under five years old participated in a structured intervention integrating Smart Discrete Trial Training and Smart Discrimination Training across five tasks focused on uppercase vowel recognition (A, O, U, E, I). The mastery criterion was set at ≥80% accuracy for three consecutive sessions. Results showed that all participants achieved reading mastery within an average of 46 sessions (approximately two weeks) and demonstrated consistent retention during generalization. Smart ABA improved reading acquisition, attention, and compliance, and provided measurable evidence of learning stability. Beyond empirical outcomes, Smart ABA contributes methodologically through explicit time-based tracking of mastery and aligns with Sustainable Development Goal (SDG) 4 by promoting inclusive and quality education for all children.

Keywords:

Autism Spectrum Disorder; Smart ABA; Reading uppercase vowels; Therapist reliability and assessment; SDGs

Introduction

Autism spectrum disorder (ASD) is a severe neurodevelopmental disorder. It often appears in early childhood and is characterized by core symptoms: repetitive, stereotyped behavior patterns and impairments in social communication and social interaction (American Psychiatric Association, 2022). In addition, children with Autism Spectrum Disorder (ASD) frequently encounter significant barriers during the early stages of reading acquisition. These challenges include difficulties in visual letter discrimination, letter-sound association, and maintaining focus during structured learning activities (Arciuli & Bailey, 2021). Although some children are capable of recognizing basic letters, many fail to consistently associate letters with their corresponding phonemes (Solis et al., 2016). Reading performance also tends to decline when



www.iejee.com
 ISSN: 1307-9298

2026 Published by KURA Education & Publishing.
 This is an open access article under the CC BY-NC-ND license. (<https://creativecommons.org/licenses/by/4.0/>)

learning stimuli change or become more complex (Randi et al., 2010). Moreover, deficits in attention and weak stimulus control are strongly correlated with low literacy performance among children with ASD (Henderson et al., 2022; Ke et al., 2023; Tsuji et al., 2024; Westerveld et al., 2021). Therefore, an instructional approach that integrates systematic structure, effective teaching techniques, and quantifiable, transparent measurement is essential to address these challenges.

Applied Behavior Analysis (ABA) has long been recognized as an effective therapeutic and instructional method for children with ASD. ABA emphasizes structured behavioral principles such as reinforcement, stimulus control, and repeated practice to support skill acquisition and behavioral adaptation. These principles have been successfully applied to improve various domains, including communication, social interaction, and academic readiness (Coburn et al., 2022; Du et al., 2024; Makrygianni et al., 2023; Sandbank et al., 2024). However, the application of ABA in reading instruction for children with ASD is still not consistently implemented through structured and measurable procedures (Arciuli & Bailey, 2021; Whalon et al., 2009). As a result, the integration of phonemic awareness, decoding, and fluency often varies across intervention settings.

Empirical evidence supports the effectiveness of Applied Behavior Analysis (ABA) in improving reading-related skills among children with ASD. Singh et al. (2017) confirmed the effectiveness of Behavior Skills Training in enhancing reading comprehension up to 80% after 12 sessions, while Solis et al. (2016) reported comprehension gains of 73-83% following a two-week ABA-based intervention. Bekmurat and Autayeva (2023) found significant lexical improvement after nine months of adapted ABA therapy that combined discrete-trial and functional communication training. Reisener et al. (2014) observed marked increases in oral reading fluency after 30 sessions using stimulus-control-based procedures such as Repeated Reading. These findings indicate that ABA can support reading development. However, several important limitations remain. Most studies do not clearly describe systematic teaching procedures, standardized instructional sequences, or the duration required to achieve mastery. In addition, the learning process is often not reported in a way that allows replication across settings.

Addressing this methodological gap, Smart Applied Behavior Analysis (Smart ABA) was developed as an extension of ABA (Anwar et al., 2022). It integrates systematic procedures, consistent teaching methods, and transparent quantitative measurement to define clear learning stages and timelines for reading mastery, as described in previous Smart ABA studies (Arneliza et

al., 2026). Smart ABA also introduces a structured two-on-one instructional model to improve consistency and engagement during instruction. This study aims to examine the application and effectiveness of Smart ABA in teaching early reading skills to children with ASD. It also explores its contribution to the Sustainable Development Goals (SDGs), particularly Goal 4, which promotes quality and inclusive education.

The study focuses on early reading as a foundational literacy skill. While prior studies have established the general efficacy of ABA in enhancing academic performance, few studies have provided detailed and measurable instructional procedures for early reading in children with ASD. In contrast, Smart ABA offers a systematic instructional framework, consistent teaching practices across therapists, and quantifiable, transparent evaluation metrics. The study also examines performance consistency across sessions and maintenance phases. This study contributes to the literature by providing a structured, measurable, and replicable instructional model for early reading. Detailed procedures are presented in the Methods section.

Accordingly, the following research questions guided this study:

1. How can a systematic and measurable reading instruction model for children with ASD be developed using Smart ABA?
2. How effective is Smart ABA in improving early reading skills among children with ASD?
3. To what extent are reading skills maintained following the implementation of Smart ABA?

Method

Research Design

This study employed a developmental research approach combined with limited implementation to examine the initial effectiveness of the Smart ABA framework in teaching early vowel reading to children with Autism Spectrum Disorder (ASD). The development phase involved designing a structured instructional model based on the Smart ABA Model, which emphasizes systematic procedures, sequential learning, and transparent measurement. The model was developed through collaborative processes and refined to ensure practical applicability in real therapy settings. This approach aligns with developmental research that focuses on model design, feasibility, and initial impact on users (Richey & Klein, 2007). The effectiveness in this study is positioned as an early indicator of functional application.

Following the development phase, a Single Subject Research (SSR) A-B design with maintenance and mastery criterion was implemented to observe changes in participants' reading performance through repeated measurement (Kazdin, 1982). The study consisted of a baseline phase (A), an intervention phase (B), and a maintenance phase to assess skill stability. Baseline was defined as an initial performance documentation phase within a structured mastery-based framework, rather than a no-assistance condition. Initial prompt procedures were applied based on Smart ABA decision rules due to the absence of prerequisite skills. The intervention emphasized continuous and progressive learning, where instruction was not withdrawn for experimental purposes, in line with ethical considerations in pediatric intervention research (Lepola et al., 2024; Matsui et al., 2021).

Participants and setting

This study involved five children with ASD under the age of five (three boys and two girls). Based on medical assessments, parent reports, and the initial Smart ABA evaluation, all participants were categorized as nonverbal. Participants were selected based on the following inclusion criteria: (1) a confirmed diagnosis of Autism Spectrum Disorder (ASD), (2) classified as nonverbal, defined as the inability to produce meaningful spoken words, and (3) not receiving other concurrent behavioral interventions during the study period. The intervention began with a basic curriculum consisting of fifteen learning activities. Parents provided written informed consent to participate in the study and were allowed to observe therapy sessions through individual closed-circuit television (CCTV) monitors installed outside the therapy room.

Development and Procedures

The development of Smart ABA was carried out through a focus group discussion involving three therapists, one supervisor, one program director, and one consultant. The team collaboratively designed a systematic instructional framework to teach reading skills to children with ASD, ensuring that it was practical, accessible, and applicable for therapists, parents, and other professionals. The Smart ABA Model is also designed to ensure that learning takes place in an enjoyable and engaging atmosphere for children, supported by warm interaction, responsive communication, and enthusiastic reinforcement. The Smart ABA approach adopts an intensive model aligned with the principles of Early Intensive Behavioral Intervention (Matson et al., 2012; Sandbank et al., 2024). Session frequency was age adjusted: four sessions per day for children under five years old, five sessions for ages 5-10, and six sessions for children above 10. All sessions were conducted by certified Smart ABA therapists trained in structured literacy-

based behavioral instruction (Arneliza et al., 2025).

Unlike conventional ABA's one-to-one format (Gelkop et al., 2015; Wood et al., 2021), Smart ABA implements a two-to-one instructional model in which one participant is supported by two trained therapists. Both therapists had completed approximately three months of structured training and internship. During each session, the therapists alternated roles as therapist and therapist assistant (TA). The therapist was responsible for delivering instructions, guiding the learning process, and preparing session reports. The therapist assistant (TA) supported session preparation, delivered prompts when needed, recorded all participant responses using trial-by-trial procedures, calculated total scores, and plotted data into graphs. This structured role division was designed to maintain child engagement, ensure systematic data recording, and minimize inter-rater bias during the intervention.

The instructional framework and assessment tools were evaluated for readability by seven therapists, all of whom reported that the procedures were easy to understand. Next, four educational and psychological experts and three autism therapy practitioners rated the clarity of the content, ease of use, and appropriateness of the language using a five-point scale. The results were analyzed using Aiken's V coefficient (Aiken, 1985). The implementation followed a continuous and sequential learning process. After participants mastered the imitation of the vowel sound "A," the program progressed directly to reading the corresponding uppercase letter "A" and continued sequentially to subsequent vowel targets. In this study, after approximately 2 to 3 months from the start of the Smart ABA program, participants began to demonstrate the ability to imitate vowel sounds. Instruction was then immediately advanced to reading tasks using uppercase letter cards. This structured progression reflects the hierarchical and continuous nature of the Smart ABA learning model, in which instruction is not interrupted once a skill is acquired. All therapy sessions were conducted in an individual therapy room with a structured environment designed to minimize distractions and support consistent learning implementation.

Data Collection Techniques and Measurement

All participant responses were systematically recorded by the therapist assistant (TA) using standardized Smart Measurement and Evaluation (Smart ME) protocols (Arneliza et al., 2026). Role rotation between therapist and assistant served as an inter-rater reliability control to minimize scoring bias. Each response was coded following Smart DTT procedures (Anwar et al., 2022) and plotted in real-time graphs for continuous monitoring by parents, therapists, and consultants. Data collection was conducted consistently across both daily and maintenance programs to ensure

accuracy and continuity. Any signs of stagnation or decline were reviewed during weekly team meetings with the consultant, program director, and supervisor, who provided corrective recommendations. All data were securely stored and reviewed under ethical supervision to ensure confidentiality, data integrity, and compliance with research standards.

Internal Validity

To ensure internal validity, this study adhered to the principles outlined by Shadish, Cook, and Campbell (2002). First, all participants had a confirmed diagnosis of ASD by a certified pediatric specialist. Second, no additional interventions were introduced during the study apart from the Smart ABA program. Third, participants followed a medically supervised diet free of casein, gluten, sugar, chemical additives, and high-phenol fruits or vegetables and were not exposed to electronic devices or television. Fourth, each child was supported by three trained therapists who alternated across four daily sessions (Sunday to Friday). The involvement of qualified ABA practitioners is critical to maintaining both safety and therapeutic effectiveness (Clark & Wilder, 2024; Ragulan et al., 2023). Data recording was performed in real time by assistant therapists, and any issues were reviewed in weekly meetings with the supervisor and consultant. Fifth, all sessions were conducted in a distraction-free, soundproof therapy room equipped with CCTV and external monitors for parental observation.

Data analysis

Visual analysis is the primary method for interpreting data in Single-Subject Research (SSR) (Dowdy et al., 2022; Ninci et al., 2015). This approach facilitates the identification of behavioral change patterns and supports systematic and quantitative evaluation of the relationship between the independent variable (Smart ABA) and the dependent variable (reading ability in children with ASD) (Blair & Ledford, 2022; Lane & Gast, 2014; Wolfe et al., 2019).

Ethical Considerations

All research procedures adhered to institutional ethical standards and regulations and were approved Research Ethics Committee of Universitas Muslim Indonesia (Approval No. 589/A). by the Research Ethics Committee of Universitas (Approval No.) Written informed consent was obtained from all participants' parents prior to data collection, and all personal identities were protected in accordance with strict confidentiality principles.

Results

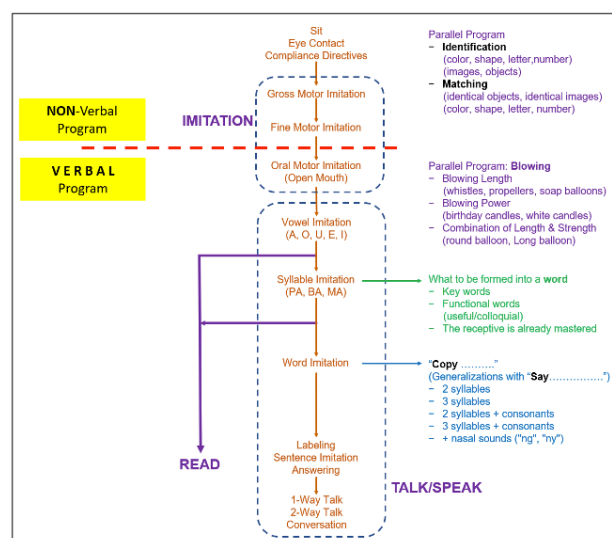
The total Aiken's V value was 0.96 ($p < 0.01$), indicating that all items were valid and the overall structure

statistically reliable. Thus, the systematic design and teaching framework for reading in children with ASD, along with its quantified assessment, were validated. The development of a reading instruction procedure for children with Autism Spectrum Disorder (ASD) resulted in a conceptual model entitled Systematic Approach to Teaching Reading in ASD Using Smart ABA, which emphasizes a structured, sequential, and quantitatively measurable learning framework. The model integrates Smart Discrete Trial Training (Smart DTT) and Smart Discrimination Training (Smart DT) (Sutadi et al., 2026) across progressive stages of reading instruction, beginning with uppercase vowel recognition (A, O, U, E, I). Each stage includes a transparent mastery criterion ($\geq 80\%$ for three consecutive sessions under a dropout system) and a time-based evaluation mechanism that enables objective and continuous monitoring of the child's developmental progress.

Model Conceptualization

The development of the systematic Teaching Speaking and Reading approach using Smart ABA is illustrated in Figure 1 below:

Figure 1.
Systematic Approach to Teaching Speaking and Reading in ASD Using Smart ABA

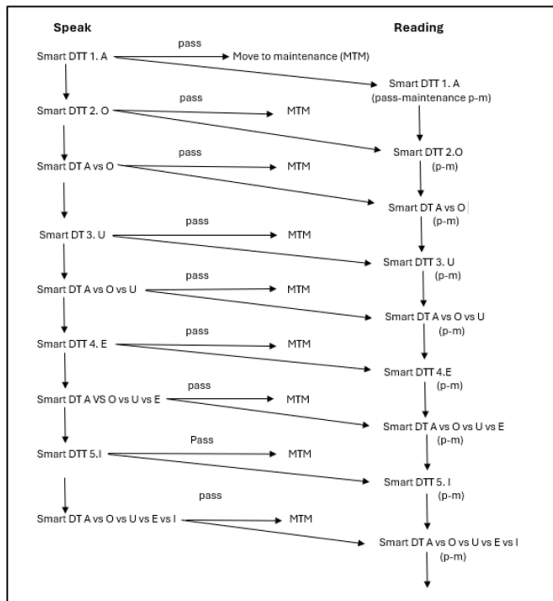


The Smart ABA stages begin with an attending-skills/compliance program, including sitting calmly, making eye contact, and following verbal instructions. Programs are organized in blocks of five activities and run in parallel with gross motor imitation exercises and 15 other programs, implemented in both daily and maintenance sessions. After mastering seven gross motor imitation activities, children proceed to fine motor imitation; after mastering seven fine motor imitation activities, they continue with oral motor imitation (only opening the mouth). In parallel, children are gradually trained to blow using various media,

including balloons, whistles, straws, paper pinwheels, and others. After completing the oral motor imitation program, daily sessions continued with Smart DTT Activity 1 (imitating the vowel 'A'). The systematics of teaching reading to children with ASD can be seen in Figure 2 below.

Figure 2.

Systematic Approach to Teaching Reading in ASD Using Smart ABA



The Smart ABA reading program is designed to be hierarchical, parallel, and measurable, following the mastery of vowel imitation skills as the main prerequisite before introducing uppercase vowel reading. This prerequisite also applies to syllables and words. Instruction is delivered through two primary formats: Smart Discrete Trial Training (Smart DTT), used to teach the imitation of a single sound or letter in each activity; and Smart Discrimination Training (Smart DT), used to train discrimination between two or more sounds or letters presented simultaneously.

The mastery criterion is defined as achieving $\geq 80\%$ accuracy across three consecutive sessions, applying a dropout system. The program consists of five main activities: reading uppercase vowels A, O, U, E, and I. The fundamental principle of Smart ABA is that children can only proceed to the reading phase once they have mastered imitation of the corresponding vowel or syllable.

Activity 1. Vowel "A"

The program begins with Smart DTT Activity 1 (imitating the sound "A"). Upon reaching the mastery criterion, the sound "A" is transferred to maintenance. In parallel, the child begins Smart DTT Reading Activity 1 (reading uppercase vowel "A"). After meeting the mastery criterion, reading activity "A" is also moved to maintenance to ensure retention.

Activity 2. Vowel "O"

The next stage is Smart DTT Activity 2 (imitating the sound "O"). Once the child achieves mastery, the sound "O" is placed in maintenance alongside "A." Smart DT (A vs. O) is then introduced to strengthen auditory discrimination. After meeting the criterion, the activity is moved to maintenance, and the previous imitation tasks ("A" and "O") in maintenance are discontinued. In the reading program, after the child masters uppercase vowel "A," they proceed to Smart DTT Reading Activity 2 ("O"). After meeting the mastery criterion, it is moved to maintenance in parallel with "A." Smart DT Reading (A vs. O) follows to train visual discrimination between the two letters. Upon mastery, reading activities "A" and "O" in maintenance are discontinued.

Activity 3. Vowel "U"

The next phase is Smart DTT Activity 3 (imitating the sound "U"). After mastery, the sound "U" is placed in maintenance, followed by Smart DT (A vs. O vs. U) to strengthen three-sound discrimination. Upon meeting the criterion, it is moved to maintenance, and the "U" imitation task in maintenance is discontinued. In parallel, the child proceeds to Smart DTT Reading Activity 3 (uppercase vowel "U"). After reaching the criterion, Smart DT Reading (A vs. O vs. U) is conducted to increase visual and auditory discrimination complexity. When mastery is achieved, the reading activity "U" is moved to maintenance, and earlier maintenance activities are discontinued.

Activity 4. Vowel "E"

This stage begins with Smart DTT Activity 4 (imitating the sound "E"). After achieving mastery, the sound "E" is moved to maintenance, followed by Smart DT (A vs. O vs. U vs. E) to train four-sound discrimination. Once the mastery criterion is met, the activity is transferred to maintenance, and the "E" imitation task in maintenance is discontinued. In parallel, the child participates in Smart DTT Reading Activity 4 (uppercase vowel "E"). After meeting the criterion, reading "E" is placed in maintenance and run concurrently with other vowels. Smart DT Reading (A vs. O vs. U vs. E) is then introduced to reinforce visual discrimination of four vowels. After mastery, reading activity "E" and the previous combinations in maintenance are discontinued.

Activity 5. Vowel "I" (Block I Integration)

The final phase involves Smart DTT Activity 5 (imitating the sound "I"). Upon meeting the mastery criterion, the sound "I" is moved to maintenance, followed by Smart DT (A vs. O vs. U vs. E vs. I) known as Smart DT Block I to integrate all vowel sounds. After mastery, the activity is transferred to maintenance, and all previous imitation

tasks are concluded. In parallel, the child undertakes Smart DTT Reading Activity 5 (uppercase vowel "I"). Once mastery is achieved, the reading activity "I" is moved to maintenance, followed by Smart DT Reading (A vs. O vs. U vs. E vs. I) or Block I Reading to ensure complete integration of auditory and visual discrimination skills. After meeting the criterion, this activity is transferred to maintenance, and all previous reading tasks are terminated.

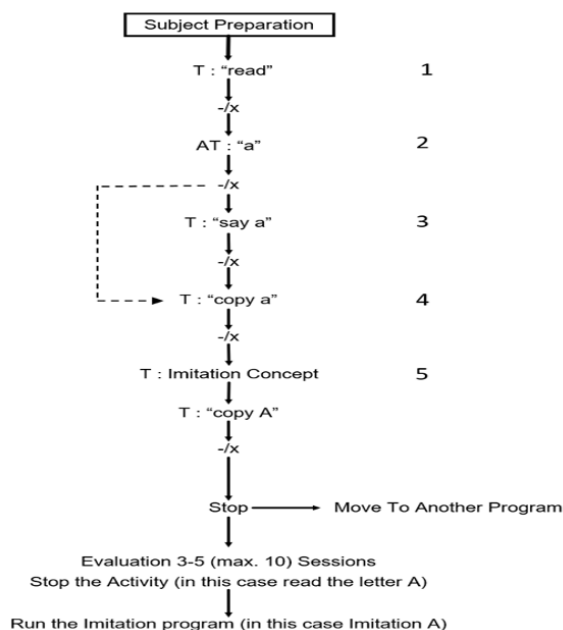
Maintenance and Generalization Phase

The maintenance program is implemented regularly to ensure retention and stability of acquired skills. The mastery criterion for the maintenance phase is defined as achieving 80% of total sessions with a score above 80%, conducted over 4–6 weeks. In this study, each participant was supported by three therapists; thus, the mastery criterion was calculated as: 3 therapists × 6 weeks = 18 sessions × 80% = 15 sessions (meaning the child passes if they achieve ≥15 out of 18 sessions with >80% accuracy). After all Smart DT combinations (both imitation and reading) reach the mastery criteria, children proceed to the generalization phase at home. Parents, who have received brief training from Smart ABA therapists, continue parallel sessions.

Escalation Prompt

In a reading program, if the child does not respond, remains silent, makes mistakes, or is off-task, the therapist and therapy assistant can implement the Smart Escalation Prompt technique (tiered prompting). The structure of the Escalation Prompt is illustrated in Figure 3 below.

Figure 3:
Standard Operating Procedure for Implementing Smart Escalation Prompt



In Figure 3, T is the therapist, and AT is the therapist's assistant. In Figure 31, it starts with preparing the child, then the therapist gives verbal program instructions by labeling (e.g. "What color?"; "What is this?") or reading (e.g. "read") or answering questions (e.g. "What is your name?"). In the example above, the program reads the letter card A with the instruction "read" with the expected response from the child, namely the child answers "a".

The requirement for this program/activity to read the letter A is that the child has been able to imitate "a". If the child does not respond correctly (in this case marked with -/x, i.e. a silent or incorrect response), the therapist's assistant sitting behind the child prompts from the back of the child's side by saying "a" in a rather quiet voice that is quite audible to the child. If the child still responds silently or incorrectly (-/x), then the prompt is taken over by the therapist, with the therapist saying "say a" if the child has previously generalized from "copy a" to "say a" in the say drill program.

However, if there has never been a generalization of "copy a" with "Say a", then the therapist immediately gives the prompt "copy a". If the child continues to respond silently or incorrectly (-/x), the therapist implements the concept of imitation. The concept of imitation is to restore the child's understanding of the concept of imitation, with the therapist giving instructions to imitate gross and/or fine motor movements, 3-5 movements that the therapist is sure the child is capable of, so that it runs smoothly, does not stutter, then the therapist again gives instructions to "imitate A" with the mention of "A" expressively. If the child still responds silently or incorrectly (-/x), then the drilling is stopped and then moved to another program in the child's individual curriculum. If the child responds off-task, the child is prepared again, sat down, and paid attention, if needed.

Next, an evaluation is conducted for 3-5 sessions, a maximum of 10 sessions, in which if there is no improvement in ability, then the program/activity of reading the letter A is stopped, to re-run the program/activity of imitation A. This shows that the child's ability to imitate A has decreased. After the child's imitation ability reaches the graduation criteria, the program is run again (the example above is the letter A card reading program). What if in the Smart EP series the child responds correctly, namely following the prompt given? The value given is P (not the +/plus sign) because the child responds correctly in a series of prompts. Next, the Smart DTT series is carried out again, which starts again from preparing the child if necessary to be given instructions to label/read/answer, which in the example above is reading the letter A card.

Measurement and Evaluation

All participant responses were meticulously recorded by the assistant therapist following standardized Smart ABA assessment procedures. The rotation of roles between therapist and assistant functioned as an inter-rater reliability control to minimize scoring bias. Each response was coded according to Smart DTT procedures (Anwar et al., 2022) and visualized in real-time graphs, allowing parents, therapists, and consultants to monitor each child's progress and promptly adjust interventions in cases of stagnation or regression. Additional observations were conducted by the supervisor and consultant, both directly in the therapy room and through OCTV recordings.

Each child's response was categorized as silent (-), incorrect (X), off-task (O), partially correct (P+), or correct (+). Correct responses produced after direct prompting from the assistant on the third instruction were coded as P. Approximate responses in the verbal imitation program were coded as A (approximation), while responses in labeling and question-answering tasks were marked with a plus sign (+). The assistant therapist calculated session scores using the formula: $\text{Score} = (\text{Correct Response} / \text{Opportunity}) \times 100\%$, where CR represents the number of correct responses and OP the number of response opportunities (excluding P). Mastery was defined as achieving $\geq 80\%$ for three consecutive sessions under a dropout system (Arneliza et al., 2026). Each session consisted of three to five instructional cycles, two-instruction, or single-instruction depending on the child's performance and consultant recommendations. An example of SmartABA recording and scoring is presented in Table 1.

Table 1.
Response Recording and Scoring

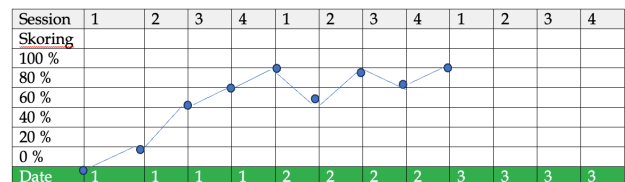
		TRIAL										TOTAL		
Date	Session	1	2	3	4	5	6	7	8	9	10	Op	CR	Score
1	1	X	X	P	X	X	P	X	X	P	X	10	0	0%
		X	P	X	X	P								
1	2	X	X	P	X	X	P	-	O	P	X	15	3	20%
		X	P	X	P	X	X	P	+	+	+			
1	3	-	X	P	+	+	P	-	+	+	P+	13	8	61%
		+	+	+	+									
1	4	+	+	O	+	+	+	+				7	6	85%
2	1	+	+	+	+	+						5	5	100%
2	2	X	+	+	X	+	+	+	+	+		9	7	77%
2	3	+	+	+	+							4	4	100%
2	4	+	+	O	+	+	+	+	+			8	7	87%
3	1	+	+	+	+	+						5	5	100%
		Fall					Passed/Achieved							

Data Recording and Monitoring

Table 1 shows that some sessions consisted of more than five cycles due to the implementation of Smart

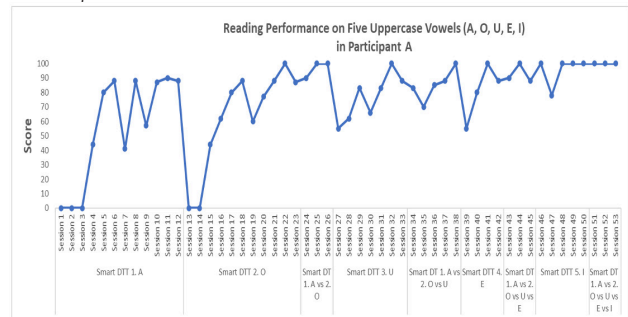
Establishing Operation (Smart EO), which lies outside the scope of this study. In each session, participants' percentage scores were plotted on progress graphs that served as visual tools for monitoring development and identifying potential regression. Recording was consistently carried out across both daily and maintenance programs. When stagnation or decline was observed, results were discussed during weekly meetings involving the consultant, program director, supervisor, and all therapists. The professional team then provided recommendations for corrective action to optimize learning outcomes. An example of the graphical recording format is presented in Table 2.

Table 2.
Graph of scoring calculation



The instrument was then implemented with five ASD participants by fifteen certified Smart ABA therapists. The results of implementing the systematic procedure for teaching reading in children with ASD using Smart ABA for Participant A are presented in Figure 4.

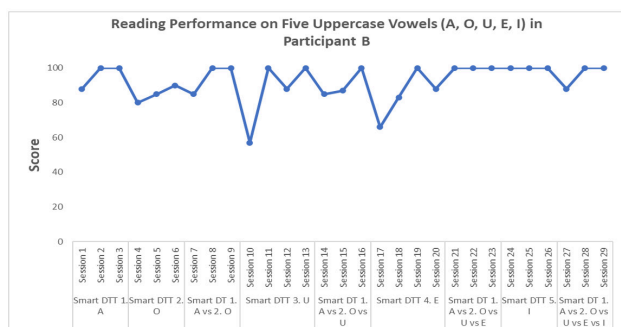
Figure 4:
Reading Performance on Uppercase Vowels in Participant A



The first two activities (A and O) were the most challenging for Participant A. Figure 4 shows that a longer time was required to reach the mastery criterion at this stage. The responses were still unstable and frequently declined. This indicates that the initial learning phase required greater effort and adjustment. However, after passing this stage, the learning process became more efficient. In subsequent letters, Participant A showed more consistent improvement. The time needed to reach mastery gradually decreased. By the end of the program, performance became stable and was well maintained across sessions.

The results of implementing the systematic procedure for teaching reading in children with ASD using Smart ABA for Participant B are presented in Figure 5.

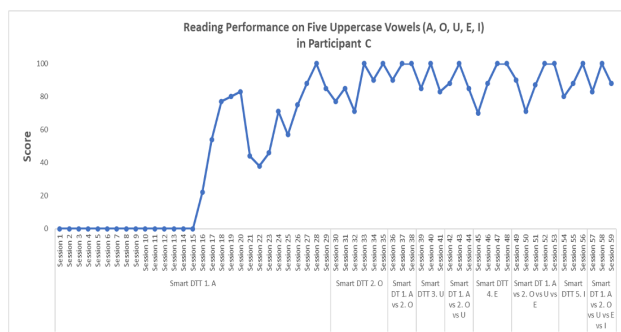
Figure 5:
Reading Performance on Uppercase Vowels in Participant B



Participant B showed generally high performance across sessions, with scores frequently reaching 90–100. Early sessions were relatively stable, although occasional declines appeared, particularly around Sessions 10 and 17. These drops indicate temporary difficulty when new targets were introduced. However, recovery was quick, and performance improved again in subsequent sessions. As training progressed, responses became more consistent. By the later sessions, Participant B maintained stable and high scores, suggesting effective learning and good retention of the vowel reading skills.

The results of implementing the systematic procedure for teaching reading in children with ASD using Smart ABA for Participant C are presented in Figure 6.

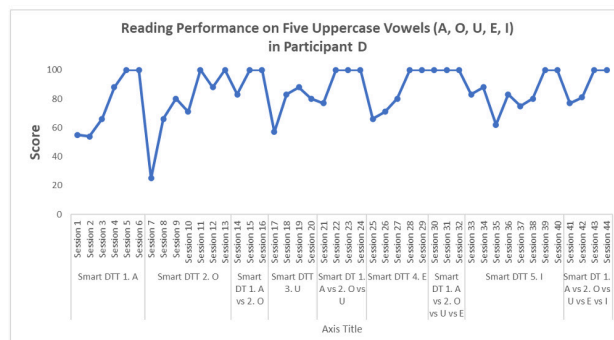
Figure 6:
Reading Performance on Uppercase Vowels in Participant C



Participant C showed very low performance in the early sessions, with scores remaining at zero until around Session 15. A sharp increase appeared after Session 15, indicating the initial emergence of the skill. However, performance was still unstable, with noticeable fluctuations in the middle phase. Several drops occurred when new targets were introduced. Despite this, improvement continued. In later sessions, scores frequently reached 90–100 and became more consistent. This pattern indicates gradual learning, improved adaptation, and stable retention over time.

The results of implementing the systematic procedure for teaching reading in children with ASD using Smart ABA for Participant D are presented in Figure 7.

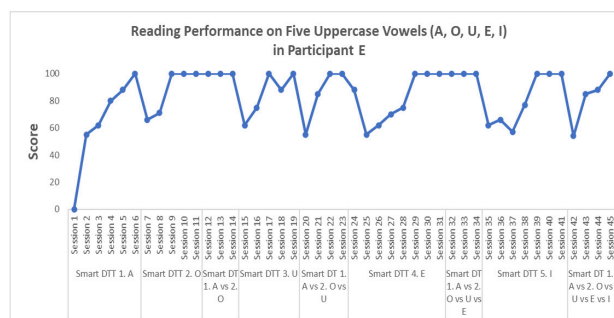
Figure 7:
Reading Performance on Uppercase Vowels in Participant D



Participant D demonstrated relatively high performance from the initial sessions. Scores rose to high levels within a short period. However, several sharp dips were observed, particularly around Sessions 7, 17, and 35, indicating temporary difficulties or adjustments. Although fluctuating, recovery was rapid after the dips. In subsequent sessions, performance returned to high levels and became more stable. By the end of the program, scores were consistently high, indicating strong learning, good adaptation, and stable retention of vowel reading skills.

The results of implementing the systematic procedure for teaching reading in children with ASD using Smart ABA for Participant E are presented in Figure 8.

Figure 8:
Reading Performance on Uppercase Vowels in Participant E



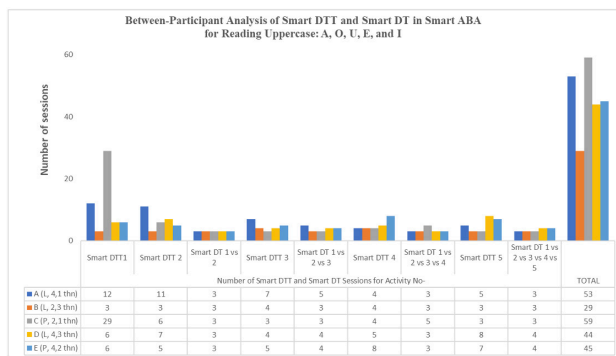
Participant E showed rapid performance improvement after the initial sessions. Scores rose rapidly, from very low to high levels. This indicates early mastery of the skill, although there were some fluctuations throughout the sessions, particularly when new targets appeared. Some dips occurred, but recovery was relatively rapid. In many sessions, scores reached 100, indicating strong performance. Toward the end, responses became more consistent. Overall, this pattern indicates effective learning, good adaptation,

and stable retention of vowel reading skills despite occasional variability.

To further clarify these individual differences, Figure 9 provides a between-participant comparison of session distribution across instructional phases, highlighting variations in learning efficiency and progression under the Smart ABA framework.

Figure 9.

Results of reading five uppercase vowel letters (A, O, U, E, I) from five participants



All five participants showed variations in the time required to achieve mastery of five uppercase vowel letters (A, O, U, E, I) using the Smart ABA framework. Participant A (4 years, 1 month) completed the program in 53 sessions, while Participant B (2 years, 3 months) achieved mastery fastest in 29 sessions. Participant C (2 years, 1 month) required 59 sessions, whereas Participants D (4 years, 3 months) and E (4 years, 2 months) completed their programs in 44 and 45 sessions, respectively. The overall average was 46 sessions, equivalent to approximately two weeks, demonstrating the efficiency of Smart ABA in accelerating reading acquisition with consistent performance across all participants.

The initial Smart DTT phase proved to be the most challenging, as it marked the children's first exposure to reading. Subsequent Smart DT 1 vs 2 phases demanded higher focus and behavioral control due to increasing discriminative complexity. Following these stages, the number of required sessions gradually declined, reflecting improvements in compliance, attention, and cognitive engagement. This trend indicates that Smart ABA not only enhances academic skills but also effectively stimulates adaptive behavior, focused attention, and discrimination abilities through its systematic and structured learning process.

Comparisons among participants revealed that chronological age was not the primary determinant of learning speed. Variations were more strongly associated with behavioral readiness, instructional compliance, and reinforcement consistency throughout the learning process. All participants reached reading mastery within an average of two

weeks (11–12 effective days), while home-based generalization showed 100% retention after three days. These findings confirm Smart ABA's effectiveness as a structured, measurable, and replicable instructional model that supports early literacy development in children with ASD, both in therapeutic and home-based learning contexts.

Discussion

Autism Spectrum Disorder (ASD) is a neurodevelopmental condition marked by persistent deficits in communication, social interaction, and adaptive functioning, typically appearing before the age of three (Blaxill et al., 2022; Salari et al., 2022). The global prevalence of ASD has risen significantly over the past two decades, underscoring the urgency of early, targeted interventions to improve developmental outcomes (Jiang et al., 2024). One of the most challenging areas for children with ASD is verbal communication, which affects their capacity for social participation, learning, and independence (Dehqanova & Tagonova, 2022; Tawil et al., 2018). Reading requires the integration of visual, linguistic, and cognitive processes, domains often impaired in autism due to attentional and sensory deficits (Linderholm et al., 2018; Kim & Klassen, 2018). Enhancing reading ability has been shown to improve communication, autonomy, and academic achievement (Boyle et al., 2019; Light et al., 2019). Bandura's Social Learning Theory (Bandura, 1977) highlights the role of observation and modeling in developing literacy skills among children with ASD.

Applied Behavior Analysis (ABA) offers an evidence-based framework for modifying behavior through reinforcement, prompting, and stimulus control (Du et al., 2024). Its intensive form, Early Intensive Behavioral Intervention (EIBI), provides structured and repetitive instruction that strengthens language, cognition, and social adaptation (Bull Ellegaard et al., 2024; Sandbank et al., 2024). Smart ABA advances ABA by integrating systematic instructional sequencing, a two-on-one therapist model, and real-time quantification of progress (Arneliza et al., 2026), addressing methodological gaps through transparent and replicable teaching procedures aligned with Sustainable Development Goal 4, which promotes inclusive and quality education for all. An important distinction in instructional format lies between the commonly used one-on-one ABA approach and the two-on-one model implemented in Smart ABA.

In many ABA applications, instruction is delivered by a single therapist who is also responsible for recording responses (Reichow et al., 2012; Virués-Ortega, 2010; Wong et al., 2015). While this approach is effective, it may limit real-time data accuracy and reduce opportunities for immediate observational support. In contrast, the two-on-one model in Smart ABA allows

one therapist to focus on instruction while the therapist assistant (TA) records responses systematically and provides prompt support when needed. This division of roles enhances data accuracy, maintains child engagement, and reduces inter-rater bias. The presence of two trained therapists also enables continuous monitoring and immediate adjustment of instructional strategies, which may contribute to faster skill acquisition and more stable performance outcomes observed in this study. The use of a two-on-one instructional model in Smart ABA is grounded in the need for precise and continuous behavioral monitoring. In one-on-one settings, therapists are required to simultaneously deliver instruction and record responses, which may divide attention and reduce the accuracy of real-time data capture. In contrast, the two-on-one model allows a clear division of roles, where one therapist focuses on instruction while the therapist assistant (TA) ensures accurate trial-by-trial recording and immediate prompt delivery. This structure supports more precise data collection, faster instructional adjustments, and reduced response latency, which are critical in early skill acquisition, particularly in children with ASD who require consistent and immediate feedback.

This study suggests that Smart ABA may be more effective in supporting early reading development among children with ASD. While ABA has been widely applied to teach reading-related skills, previous studies have not consistently provided explicit descriptions of instructional sequencing, mastery timelines, or systematic measurement procedures. In contrast, Smart ABA integrates structured instructional sequencing, a two-on-one model (one child with two therapists), and real-time quantitative measurement to support objective progress evaluation. The findings indicate that this approach may facilitate faster skill acquisition and enhance performance stability across intervention phases. In this study, participants achieved the mastery criterion for five uppercase vowels (A, O, U, E, and I) in an average of 46 sessions, or approximately two weeks. The use of a clearly defined mastery criterion allows progression to subsequent learning stages to be determined objectively, ensuring that each skill is acquired consistently before advancing. This duration appears more efficient compared to reports from previous ABA-based studies that did not explicitly document mastery timelines. These results suggest that the structured and measurable nature of Smart ABA may contribute to more efficient and consistent learning outcomes.

The findings also revealed progressive improvements in compliance, attention, and cognitive ability, suggesting that the Smart ABA system may support adaptive behavior and early literacy through systematic, structured, and measurable procedures. Singh et al. (2017) reported an 80% improvement in

reading comprehension after 12 sessions, Solis et al. (2016) achieved 73–83% gains over two weeks, and Bekmurat and Autayeva (2023) observed significant lexical improvements after nine months of adapted ABA therapy. Arciuli and Bailey (2021) highlighted the importance of quantitative monitoring in reading programs. Beckerson et al. (2024) found comprehension improvements after 200 hours of Visualizing and Verbalizing training, Reisener et al. (2014) noted fluency gains after 30 sessions, and Boyle et al. (2019b) confirmed the effectiveness of shared reading for comprehension. However, most of these studies did not clearly describe standardized instructional procedures, systematic measurement approaches, or the time required to reach the mastery criterion. Smart ABA bridges this methodological gap by providing a structured, quantifiable, and replicable instructional framework including clearly defined mastery criteria applicable across educational and therapeutic contexts. This approach may contribute to more efficient and consistent learning outcomes.

Implications for Practice

This study explicitly documents the mastery of five uppercase vowels (A, O, U, E, I) within an average of 46 sessions, or approximately two weeks, suggesting both efficiency and methodological consistency within the Smart ABA framework. The findings enrich the ABA literature by introducing time-based measurement as a meaningful metric for assessing intervention effectiveness. This approach addresses a methodological gap in ABA research by emphasizing that time tracking enhances clinical validity, data transparency, and replicability. The structured Smart ABA model provides an evidence-based instructional framework that enables objective progress monitoring, supports individualized learning, and aligns with Sustainable Development Goal 4 which aims to ensure quality and inclusive education for all children through measurable and replicable therapeutic outcomes.

Implications for Future Research

Documenting the duration required for skill mastery serves not only as an indicator of efficiency but also as an early diagnostic tool for identifying stagnation or regression. Boyle et al. (2019) highlighted the importance of structured and systematic instructional approaches in improving early language and literacy outcomes in children with ASD, which is consistent with the emphasis on systematic measurement and transparent data recording in the Smart ABA system. Future studies should extend this framework to broader literacy domains, adaptive behaviors, and diverse age groups, as well as examine its scalability within inclusive educational systems aligned with the Sustainable Development Goals (SDGs).

Conclusion

This study demonstrates that Smart ABA significantly enhances early reading skills in children with ASD through a systematic learning structure, standardized teaching techniques, a two-on-one model (one child with two therapists), and transparent quantitative measurement. Compared to conventional ABA, Smart ABA requires a shorter duration, with an average of 46 sessions or approximately two weeks to achieve mastery of five uppercase vowel letters (A, O, U, E, I), indicating greater efficiency and consistency.

Beyond its empirical impact, Smart ABA provides an important methodological contribution by introducing time-based tracking as an indicator of intervention effectiveness. These findings make a significant contribution to the development of evidence-based interventions for children with ASD and support the achievement of Sustainable Development Goal (SDG) 4, which aims to ensure inclusive and quality education for all children.

This study's main limitation lies in its small sample size of five nonverbal children with ASD under five years old, which restricts the generalizability of the results. Future research should include a larger and more diverse participant sample to enhance external validity, and investigate the effectiveness of Smart ABA across a wider range of developmental stages and literacy domains in older children with ASD.

Use of Artificial Intelligence Tools

The authors used ChatGPT as a supporting tool for language refinement, grammar checking, and improving the readability of the manuscript. All research activities, interpretation, and final decisions regarding the manuscript were made by the authors, who bear full responsibility for its content.

References

- Aiken, L. R. (1985). Three Coefficients for Analyzing the Reliability and Validity of Ratings. *Educational and Psychological Measurement*, 45(1), 131–142. <https://doi.org/10.1177/0013164485451012>
- American Psychiatric Association. (2022). *Diagnostic and Statistical Manual of Mental Disorders (5th ed., text rev.)*. American Psychiatric Association.
- Anwar, A., Sutadi, R., & Miranda, C. (2022). Development of Discrete Trial Training (DTT) Procedure in Smart Applied Behavior Analysis (Smart ABA) for Autism. *Journal of Psychology and Behavior Studies*, 2(1), 44–63. <https://doi.org/10.32996/jpbs.2022.2.1.7>
- Arciuli, J., & Bailey, B. (2021a). The promise of comprehensive early reading instruction for children with autism and recommendations for future directions. *Language, Speech, and Hearing Services in Schools*, 52(1), 225–238. https://doi.org/10.1044/2020_LSHSS-20-00019
- Arciuli, J., & Bailey, B. (2021b). The Promise of Comprehensive Early Reading Instruction for Children With Autism and Recommendations for Future Directions. *Language, Speech, and Hearing Services in Schools*, 52(1), 225–238. https://doi.org/10.1044/2020_LSHSS-20-00019
- Arneliza, A., Rahmawati, S., Sutadi, R., Alsa, A., & Yunanto, K. T. (2025). Experiences and Strategies of Competent Smart ABA Therapists in Treating Children with Autism. *Pakistan Journal of Psychological Research*, 40(2), 351–372. <https://doi.org/10.33824/PJPR.2025.40.2.21>
- Arneliza, A., Sutadi, R., Alsa, A., & Yunanto, K. T. (2026). Development and Implementation of Smart Measurement and Evaluation (Smart ME): A Trial-by-Trial Monitoring System for Autism Therapy. *Journal of Culture and Values in Education*, 9(1), 106–124. <https://doi.org/10.46303/jcve.2026.5>
- Bandura, A. (1977). *Social Learning Theory*. Prentice-Hall.
- Beckerson, M., Paisley, C., Murdaugh, D., Holm, H., Lemelman, A., Spencer, A., O'Kelley, S., & Kana, R. (2024). Reading comprehension improvement in autism. *Frontiers in Psychiatry*, 15(March), 1–13. <https://doi.org/10.3389/fpsy.2024.1292018>
- Bekmurat, A. T., & Autayeva, A. N. (2023). The Application of Adapted Applied Behavior Analysis Therapy for Developing Lexical and Semantic Skills in Preschool Children with Autism Spectrum Disorder. *European Journal of Educational Research*, 12(4), 1657–1665.
- Blair, B. J., & Ledford, J. R. (2022). *Visual Analysis in Single-Subject Designs BT - The SAGE Encyclopedia of Research Design* (B. B. Frey (ed.)). SAGE Publications. <https://doi.org/10.4135/9781071812082.n669>
- Blaxill, M., Rogers, T., & Nevison, C. (2022). RETRACTED ARTICLE: Autism Tsunami: the Impact of Rising Prevalence on the Societal Cost of Autism in the United States. *Journal of Autism and Developmental Disorders*, 52(6), 2627–2643. <https://doi.org/10.1007/s10803-021-05120-7>

- Boyle, S. A., McNaughton, D., & Chapin, S. E. (2019a). Effects of Shared Reading on the Early Language and Literacy Skills of Children With Autism Spectrum Disorders: A Systematic Review. *Focus on Autism and Other Developmental Disabilities*, 34(4), 205–214. <https://doi.org/10.1177/1088357619838276>
- Boyle, S. A., McNaughton, D., & Chapin, S. E. (2019b). Effects of Shared Reading on the Early Language and Literacy Skills of Children With Autism Spectrum Disorders: A Systematic Review. *Focus on Autism and Other Developmental Disabilities*, 34(4), 205–214. <https://doi.org/10.1177/1088357619838276>
- Bull Ellegaard, S. E., Isaksen, J., & Eldevik, S. (2024). Supervision of Early Intensive Behavioral Intervention Onsite or via Videoconference; Outcomes in a Randomized Controlled Trial Pilot. *International Electronic Journal of Elementary Education*, 16(3), 409–416. <https://iejee.com/index.php/IEJEE/article/view/2183>
- Clark, R. J., & Wilder, D. A. (2024). An evaluation of training components necessary to teach staff members to conduct an intensive pediatric feeding intervention. *Behavioral Interventions*, 39(3). <https://doi.org/10.1002/bin.2004>
- Coburn, K. L., Kurtz, M. R., Rivera, D., & Kana, R. K. (2022). Behavioral and neurobiological evidence for the effects of reading interventions on autistic children: A systematic review. *Neuroscience & Biobehavioral Reviews*, 139, 104748. <https://doi.org/10.1016/j.neubiorev.2022.104748>
- Dehqonova, M., & Tagonova, G. (2022). The Importance of Didactic Games in Speech Therapy in the Development of Speech in Children with Autism and the Ability to Choose Effective Methods. *European Journal of Innovation In Nonformal Education*, 2(1), 133–137.
- Dowdy, A., Jessel, J., Saini, V., & Peltier, C. (2022). Structured visual analysis of single-case experimental design data: Developments and technological advancements. *Journal of Applied Behavior Analysis*, 55(2), 451–462. <https://doi.org/https://doi.org/10.1002/jaba.899>
- Du, G., Guo, Y., & Xu, W. (2024a). The effectiveness of applied behavior analysis program training on enhancing autistic children's emotional-social skills. *BMC Psychology*, 12(1). <https://doi.org/10.1186/s40359-024-02045-5>
- Du, G., Guo, Y., & Xu, W. (2024b). The effectiveness of applied behavior analysis program training on enhancing autistic children ' s emotional-social skills. *BMC Psychology*.
- Gelkop, N., Burshtein, D. G., Lahav, A., Brezner, A., Al-Oraibi, S., Ferre, C. L., & Gordon, A. M. (2015). Efficacy of constraint-induced movement therapy and bimanual training in children with hemiplegic cerebral palsy in an educational setting. *Physical and Occupational Therapy in Pediatrics*, 35(1), 24–39. <https://doi.org/10.3109/01942638.2014.925027>
- Henderson, L. M., Clarke, P. J., & Snowling, M. J. (2022). Neural and cognitive mechanisms of attention and reading in autism spectrum disorder. *Autism Research*, 15(6), 1054–1068. <https://doi.org/10.1002/aur.2739>
- Jiang, X., Chen, X., Su, J., & Liu, N. (2024). Prevalence of autism spectrum disorder in mainland china over the past 6 years: a systematic review and meta-analysis. *BMC Psychiatry*, 24(1), 1–15. <https://doi.org/10.1186/s12888-024-05729-9>
- Ke, S., Xu, H., & Chan, R. C. K. (2023). Deficits in attentional control and their impact on literacy learning in children with autism spectrum disorder. *Research in Developmental Disabilities*, 133, 104398. <https://doi.org/10.1016/j.ridd.2023.104398>
- Kim, L. E., & Klassen, R. M. (2018). Teachers' cognitive processing of complex school-based scenarios: Differences across experience levels. *Teaching and Teacher Education*, 73, 215–226. <https://doi.org/10.1016/j.tate.2018.04.006>
- Lane, J. D., & Gast, D. L. (2014). Visual analysis in single case experimental design studies: brief review and guidelines. *Neuropsychological Rehabilitation*, 24(3–4), 445–463. <https://doi.org/10.1080/09602011.2013.815636>
- Lepola, P., Nelson, R., & Matsui, K. (2024). Ethical consideration in the design and conduct of pediatric clinical trials. In *Essentials of Translational Pediatric Drug Development* (pp. 421–449). Elsevier. <https://doi.org/10.1016/B978-0-323-88459-4.00017-1>
- solis, J., McNaughton, D., & Caron, J. (2019). New and emerging AAC technology supports for children with complex communication needs and their communication partners: State of the science and future research directions. *AAC: Augmentative and Alternative Communication*, 35(1), 26–41. <https://doi.org/10.1080/07434618.2018.1557251>

- Linderholm, T., Virtue, S., Tzeng, Y., & van den Broek, P. (2018). Fluctuations in the availability of information during reading: Capturing cognitive processes using the landscape model. In *Accessibility in Text and Discourse Processing: A Special Issue of Discourse Processes* (pp. 165–186). <https://doi.org/10.4324/9781315046105-5>
- Makrygianni, M. K., Gena, A., Katoudi, S., & Galanis, P. (2023). The effectiveness of applied behavior analytic interventions for children with autism spectrum disorder: A meta-analytic study. *BMC Psychology*, 11(1), 57. <https://doi.org/10.1186/s40359-023-01150-7>
- Matson, J. L., Tureck, K., Turygin, N., Beighley, J., & Rieske, R. (2012). Trends and topics in Early Intensive Behavioral Interventions for toddlers with autism. *Research in Autism Spectrum Disorders*, 6(4), 1412–1417. <https://doi.org/10.1016/j.rasd.2012.02.010>
- Matsui, K., Ibuki, T., Tashiro, S., & Nakamura, H. (2021). Principles of ethical consideration required for clinical research involving children. *Pediatrics International*, 63(3), 248–259. <https://doi.org/10.1111/ped.14529>
- Ninoi, J., Vannest, K. J., Willson, V., & Zhang, N. (2015). Interrater Agreement Between Visual Analysts of Single-Case Data: A Meta-Analysis. *Behavior Modification*, 39(4), 510–541. <https://doi.org/10.1177/0145445515581327>
- Ragulan, S., Bertoli, E., Shinall, J. A., & Kahng, S. (2023). Effects of Acceptance and Commitment Training on Treatment Integrity Amongst Behavioral Technicians. *Behavior Modification*, 47(3), 719–751. <https://doi.org/10.1177/01454455221137327>
- Randi, J., Newman, T., & Grigorenko, E. L. (2010). Teaching Children with Autism to Read for Meaning: Challenges and Possibilities. *Journal of Autism and Developmental Disorders*, 40(7), 890–902. <https://doi.org/10.1007/s10803-010-0938-6>
- Reichow, B., Barton, E. E., Boyd, B. A., & Hume, K. (2012). Early intensive behavioral intervention (EIBI) for young children with autism spectrum disorders (ASD). *Cochrane Database of Systematic Reviews*, 10. <https://doi.org/10.1002/14651858.CD009260.pub2>
- Reisener, C. D., Lancaster, A. L., McMullin, W. A., & Ho, T. (2014). A Preliminary Investigation of Evidence-Based Interventions to Increase Oral Reading Fluency in Children With Autism. *Journal of Applied School Psychology*, 30(1), 50–67. <https://doi.org/10.1080/15377903.2013.869785>
- Richey, R. C., & Klein, J. D. (2007). *Design and Development Research: Methods, Strategies, and Issues*. Lawrence Erlbaum Associates.
- Salari, N., Rasoulpoor, S., Rasoulpoor, S., Shohaimi, S., Jafarpour, S., Abdoli, N., Khaledi-Paveh, B., & Mohammadi, M. (2022). The global prevalence of autism spectrum disorder: a comprehensive systematic review and meta-analysis. *Italian Journal of Pediatrics*, 48(1). <https://doi.org/10.1186/s13052-022-01310-w>
- Sandbank, M., Pustejovsky, J. E., Bottema-Beutel, K., Caldwell, N., Feldman, J. I., Crowley LaPoint, S., & Woynaroski, T. (2024). Determining Associations Between Intervention Amount and Outcomes for Young Autistic Children: A Meta-Analysis. *JAMA Pediatrics*, 178(8), 763–773. <https://doi.org/10.1001/jamapediatrics.2024.1832>
- Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and Quasi-experimental Designs for Generalized Causal Inference*. Houghton Mifflin.
- Singh, B. D., Moore, D. W., Furlonger, B. E., Anderson, A., Busacca, M. L., & English, D. L. (2017). Teaching Reading Comprehension Skills to a Child with Autism Using Behaviour Skills Training. *Journal of Autism and Developmental Disorders*, 47(10), 3049–3058. <https://doi.org/10.1007/s10803-017-3229-7>
- Solis, M., El Zein, F., Vaughn, S., & Falcomata, T. S. (2016). Reading comprehension interventions for students with autism spectrum disorders. *Focus on Autism and Other Developmental Disabilities*, 31(4), 266–277. <https://doi.org/10.1177/1088357615583464>
- Sutadi, R., Teguh Yunanto, K., Khattak, A. Z., Prasetyo, C., & Arneliza. (2026). Implementation of Smart Discrimination Training to Improve Discrimination Skills in Autism Spectrum Disorder. *Tazkiya Journal of Psychology*, 14(1). <https://doi.org/10.15408/tazkiya.v14i1.51035>
- Tawil, D. S., Bradford, D., & Yates, C. R. (2018). Talk to Me: The Role of Human-Robot Interaction in Improving Verbal Communication Skills in Students with Autism or Intellectual Disability. *RO-MAN 2018 - 27th IEEE International Symposium on Robot and Human Interactive Communication*, 1–6. <https://doi.org/10.1109/ROMAN.2018.8525698>

- Tsuji, H., Yamada, T., & Kato, M. (2024). Attentional modulation and reading performance in children with ASD: Behavioral and EEG evidence. *Frontiers in Psychology*, 15, 1374212. <https://doi.org/10.3389/fpsyg.2024.1374212>
- Virués-Ortega, J. (2010). Applied behavior analytic intervention for autism in early childhood: Meta-analysis, meta-regression and dose-response meta-analysis of multiple outcomes. *Clinical Psychology Review*, 30(4), 387–399. <https://doi.org/10.1016/j.cpr.2010.01.008>
- Westerveld, M. F., Paynter, J., & Trembath, D. (2021). Oral language predictors of early reading in children with autism spectrum disorder. *Journal of Autism and Developmental Disorders*, 51(3), 895–907. <https://doi.org/10.1007/s10803-020-04556-1>
- Whalon, K. J., Al Otaiba, S., & Delano, M. E. (2009). Evidence-Based Reading Instruction for Individuals With Autism Spectrum Disorders. *Focus on Autism and Other Developmental Disabilities*, 24(1), 3–16. <https://doi.org/10.1177/1088357608328515>
- Wolfe, K., Barton, E. E., & Meadan, H. (2019). Systematic Protocols for the Visual Analysis of Single-Case Research Data. *Behavior Analysis in Practice*, 12(2), 491–502. <https://doi.org/10.1007/s40617-019-00336-7>
- Wong, C., Odom, S. L., Hume, K. A., Cox, A. W., Fettig, A., Kucharczyk, S., Brock, M. E., Plavnick, J. B., Fleury, V. P., & Schultz, T. R. (2015). Evidence-Based Practices for Children, Youth, and Young Adults with Autism Spectrum Disorder: A Comprehensive Review. *Journal of Autism and Developmental Disorders*, 45(7), 1951–1966. <https://doi.org/10.1007/s10803-014-2351-z>
- Wood, J. J., Sze Wood, K., Chuen Cho, A., Rosenau, K. A., Cornejo Guevara, M., Galán, C., Bazzano, A., Zeldin, A. S., & Helleman, G. (2021). Modular cognitive behavioral therapy for autism-related symptoms in children: A randomized controlled trial. *Journal of Consulting and Clinical Psychology*, 89(2), 110. <https://doi.org/10.1037/ccp0000621>

The Relationship Between Parenting Styles and Perceived Stress Among Primary School Students: The Mediating Role of Creative and Critical Thinking Dispositions

Sevil Momeni Shabani^a, Gülten Feryal Gündüz^b, Levent Vural^{c,*}

Received : 8 April 2026
Revised : 22 June 2026
Accepted : 21 July 2026
DOI : 10.26822/iejee.2026.461

^aSevil Momeni Shabani, Department of Educational Sciences, Istanbul Kültür University, Halkalı Merkez Neighborhood, Basın Ekspres Avenue, No:11, 34303 Küçükçekmece/İstanbul- Türkiye.
E-mail: s.shabani@iku.edu.tr
ORCID: <https://orcid.org/0000-0002-3069-8281>

^bGülten Feryal Gündüz, Department of Educational Sciences, Istanbul Kültür University, Halkalı Merkez Neighborhood, Basın Ekspres Avenue, No:11, 34303 Küçükçekmece/İstanbul- Türkiye.
E-mail: g.gunduz@iku.edu.tr
ORCID: <https://orcid.org/0000-0001-8096-5598>

^c**Corresponding Author:** Levent Vural, Department of Educational Sciences, Trakya University, İsmail Hakkı Tonguç Campus, 22030 Edirne- Türkiye.
E-mail: leventvural@trakya.edu.tr
ORCID: <https://orcid.org/0000-0001-9302-6143>

Abstract

This study aimed to examine the relationship between parenting styles and perceived stress among primary school students and to determine the mediating roles of creative and critical thinking dispositions in this relationship. The study was conducted using a quantitative research approach and employed a relational survey model. The sample consisted of 598 primary school students and their parents from the European side of Istanbul. Data were collected using the Creative Thinking Scale, Critical Thinking Scale, Perceived Stress Scale, and Multidimensional Parenting Styles Assessment Scale. Descriptive statistics, independent samples t-tests, ANOVA, and mediation analyses using the PROCESS macro were conducted for data analysis. The findings indicated that students' perceived stress levels were at a lower-moderate level, whereas creative and critical thinking levels were high. The majority of parents demonstrated a positive reinforcement parenting style. Parenting styles had a limited but statistically significant effect on perceived stress. However, creative and critical thinking dispositions did not play a significant mediating role in the relationship between parenting styles and perceived stress. Overall, the findings suggest that the relationship between parenting styles and perceived stress is limited and that cognitive processes such as creative and critical thinking do not significantly mediate this relationship among primary school students.

Keywords:

Parenting Styles, Perceived Stress, Critical Thinking, Creative Thinking



www.iejee.com
ISSN: 1307-9298

2026 Published by KURA Education & Publishing.
This is an open access article under the CC BY-NC-ND license. (<https://creativecommons.org/licenses/by/4.0/>)

Introduction

Childhood is a critical developmental period during which the foundations of an individual's cognitive, emotional, and social development are established. The healthy progression of this developmental process largely depends on the environmental conditions in which the child is situated. In this context, one of the most important environmental factors influencing a child's development

is the family environment, particularly parental attitudes. The attitudes and behaviors exhibited within the family directly determine the child's personality development, social adjustment, and psychological well-being (Grolnick & Pomerantz, 2009; Yan et al., 2024). Indeed, the family is considered the most fundamental immediate environment supporting a child's development, and especially during the primary school years, parental involvement serves as a significant determinant of children's academic achievement, motivation, and school adjustment (Halimah, 2024). When evaluated within the framework of Bronfenbrenner's ecological systems theory, this situation reveals that child development is shaped through multilayered environmental interactions, with parental attitudes playing a central role at the microsystem level (Bronfenbrenner, 1979). Accordingly, parental attitudes have a significant impact on child development and adjustment outcomes.

In the literature, parental attitudes are generally categorized into three main types: democratic (authoritative), authoritarian, and permissive (Fonseca et al., 2020). Democratic parenting provides a supportive structure that allows autonomy. In contrast, authoritarian and overly controlling attitudes are reported to be negatively associated with children's psychological adjustment (Özdemir & Yalçın, 2021; Wolfradt et al., 2003). Relevant research indicates that warm, accepting, and supportive parental attitudes are positively associated with children's social and emotional development (Sümer et al., 2010; Qian et al., 2024). However, it has increasingly been emphasized in recent years that the effects of parental attitudes are not limited to emotional and behavioral outcomes but also involve systematic relationships with higher-order cognitive processes. At this point, the relationship between parental attitudes and cognitive processes has emerged as a distinct area of investigation in the literature. Findings indicate that parenting styles are significantly associated with children's academic achievement and higher-order thinking skills (Deng et al., 2023; Kaya et al., 2012). The study by Deng et al. (2023) demonstrates that parental attitudes are related to children's critical thinking skills and dispositions. Similarly, parenting approaches that support autonomy are reported to be associated with children's intrinsic motivation, self-regulation, and independent thinking processes (Soenens & Vansteenkiste, 2010). However, the majority of existing studies explain the effects of parenting attitudes on children's psychological and cognitive outcomes primarily through direct relationships (Soenens & Vansteenkiste, 2010; Ryan & Deci, 2017; Kaya et al., 2012; Deng et al., 2023). Studies investigating the cognitive mechanisms through which parenting attitudes influence children's experiences of stress appear to be limited. This highlights the need to examine the effects

of parenting attitudes on child development from a more multidimensional perspective.

Perceived stress is defined as a subjective experience that arises from an individual's evaluation of environmental demands in relation to their coping capacity (Lazarus & Folkman, 1984; as cited in Folkman, 2013). Perceived stress during childhood has significant effects on academic achievement, social relationships, and psychological adjustment (Compas et al., 2001). It is suggested that parental attitudes displayed within the family environment are among the primary determinants shaping children's perceptions of stress. Supportive parental attitudes are positively associated with individuals' coping skills and psychological well-being, whereas oppressive attitudes are reported to be positively associated with stress and anxiety levels (McLeod et al., 2007; Moran et al., 2018). However, empirical evidence regarding the cognitive processes through which this relationship emerges is limited. This highlights the need to examine potential mediating variables. During the primary school years, children begin to develop higher-order cognitive skills such as problem solving, generating alternative ideas, and evaluating situations from different perspectives. In this process, supportive parenting attitudes may foster children's abilities to think independently, generate alternative solutions, and evaluate life events more realistically, whereas authoritarian and overly controlling attitudes may hinder the development of these skills. Therefore, creative and critical thinking dispositions are thought to function as balancing cognitive mechanisms in the relationship between parenting attitudes and perceived stress by influencing how children perceive and interpret stressful situations.

From a theoretical perspective, the conceptual model of the present study is grounded in several interrelated theoretical approaches. Social Cognitive Theory emphasizes that individuals' behaviors and emotional experiences are shaped through the reciprocal interaction of personal, environmental, and cognitive factors (Bandura, 1986). This perspective suggests that parenting attitudes influence not only children's behaviors but also the ways in which they perceive and interpret environmental events. Similarly, Lev Vygotsky's (1978) sociocultural approach proposes that cognitive development is constructed through social interactions and environmental support, thereby explaining the critical role of parent-child interactions in the development of higher-order cognitive processes. At this point, Self-Determination Theory states that when parents support children's basic psychological needs—such as autonomy, competence, and relatedness—children's psychological resilience and internal resources are strengthened (Ryan & Deci, 2017). Such a supportive family environment may contribute to the

development of children's higher-order cognitive skills, enabling them to evaluate stressful situations in more adaptive ways. Indeed, the Cognitive Appraisal Theory of Stress also emphasizes that an individual's level of stress is closely related to how events are interpreted and cognitively appraised (Lazarus & Folkman, 1984; as cited in Folkman, 2013). Accordingly, the strengthening of higher-order cognitive processes may help children make more effective sense of environmental stressors and enhance their capacity to cope with stress.

Higher-order cognitive skills encompass complex mental processes such as analyzing, evaluating, and adapting information to new situations (Anderson & Krathwohl, 2001) and are closely related to skills such as creative and critical thinking. Creative thinking is associated with an individual's capacity to generate original and diverse ideas and is considered an indicator of creative potential through divergent thinking (Runco & Acar, 2012). Creativity is regarded not only as an individual trait but also as a dynamic process shaped by environmental factors. Research indicates that supportive and autonomy-supportive parental attitudes are positively associated with children's creativity, whereas oppressive and overly controlling attitudes may negatively affect overall creativity (Quan, 2025; Wang, 2023). Critical thinking is a higher-order cognitive skill encompassing processes such as analysis, evaluation, and decision-making. This skill is reported to be significantly and positively associated with metacognitive awareness (Bakır & Eğmir, 2022). Additionally, parental attitudes may indirectly influence critical thinking dispositions through variables such as self-esteem (Wang et al., 2020). The primary reason for considering creative and critical thinking as mediating variables in this study is that these cognitive dispositions may influence how children evaluate their environmental experiences and cope with stress. Creative thinking may help individuals generate alternative solutions and approach stressful situations more flexibly, whereas critical thinking may support the processes of analyzing situations and making more functional decisions. Considering that cognitive processes develop rapidly during the primary school years, the role of parenting attitudes in the development of these skills becomes particularly important in relation to children's stress experiences. Therefore, examining creative and critical thinking together may contribute to a more comprehensive understanding of the cognitive processes underlying children's experiences of stress.

Indeed, research shows that there are significant relationships between critical thinking skills and problem-solving abilities and that these cognitive skills interact with one another (Liu et al., 2025). Studies also indicate that creative learning environments support

cognitive skills and that there are relationships among higher-order skills such as creativity, critical thinking, and problem solving (Yurt, 2025). From this perspective, considering critical and creative thinking together helps individuals solve problems more effectively in both academic and everyday life. Moreover, such cognitive skills are not limited to problem-solving processes but also play a determining role in how individuals evaluate and make sense of the situations they encounter (Crilly, 2025). This, in turn, shapes individuals' perceptions of and responses to stressful experiences. The responses individuals give to the problems they encounter throughout their lives directly affect their stress experiences. Stress is a psychological and physiological condition that arises when individuals perceive environmental or internal demands as threatening (Karcı & Balcı-Çelik, 2024). As perceived stress increases, life satisfaction decreases, and students' academic and social functioning are negatively affected (Ekşi et al., 2020). Since perceived stress is a subjective experience, the same situation may be evaluated differently by different individuals. Additionally, a negative relationship has been identified between critical thinking and perceived stress (Arıbaş & Oğuzhan, 2024).

A review of the literature indicates that parenting styles, perceived stress, and higher-order cognitive skills have predominantly been examined in isolation (Kaya et al., 2012; Deng et al., 2023; McLeod et al., 2007; Moran et al., 2018). However, there is a notable scarcity of studies investigating the effects of parenting styles on psychological outcomes through cognitive processes (Soenens & Vansteenkiste, 2010). Although mediating variables have been considered in existing research, these studies have largely focused on other psychological constructs, such as self-esteem (Wang et al., 2020), and holistic models that simultaneously address both creative and critical thinking remain limited. Furthermore, despite research exploring the relationship between critical thinking and perceived stress (Arıbaş & Oğuzhan, 2024), there is a lack of studies that examine these variables within a multiple mediation framework.

Examining creative and critical thinking dispositions together is thought to contribute to a more holistic understanding of children's cognitive processes. Although there are studies in the literature addressing the relationship between parenting styles and perceived stress, research investigating the cognitive mechanisms through which this relationship emerges appears to be limited. In particular, the lack of studies examining creative and critical thinking dispositions as simultaneous mediating variables constitutes one of the distinctive aspects of the present study. Examining the simultaneous mediating roles of creative and critical thinking may reveal the interconnected

functioning of these cognitive dispositions and provide a more comprehensive understanding of the relationship between parenting styles and children's stress experiences. Furthermore, existing research has largely focused on adolescence and adulthood, whereas studies exploring the relationship between these cognitive processes and stress experiences among primary school children remain limited. However, the primary school period is considered a critical developmental stage during which higher-order cognitive skills, problem-solving approaches, and coping strategies begin to develop. In this context, the present study has the potential to provide important insights into how cognitive processes are associated with stress experiences during early developmental stages. Therefore, the primary aim of the present study is to examine the relationship between parenting styles and perceived stress among primary school students and to test the simultaneous mediating roles of creative and critical thinking dispositions in this relationship.

1. What are the levels of students' perceived stress, creative thinking, and critical thinking?
2. Which parenting styles do students' parents exhibit?
3. Do students' perceived stress levels differ according to their parents' parenting styles?
4. Does creative thinking disposition play a significant mediating role in the relationship between parenting styles and perceived stress?
5. Does critical thinking disposition play a significant mediating role in the relationship between parenting styles and perceived stress?
6. Do creative and critical thinking dispositions jointly play a significant mediating role in the relationship between parenting styles and perceived stress?

Method

This section of the study presents the research model, data collection instruments and their characteristics, the procedures carried out within the scope of data analysis, and the processes conducted to ensure validity and reliability. Additionally, explanations regarding the ethical considerations of the research are also provided.

Research Model

In this study, the research process includes quantitative methods during data collection and analysis. Within this framework, single and relational survey models, which are among the general survey models, were employed (Karasar, 2020). In general, survey research involves collecting data within a specific period in

order to describe the nature of existing conditions and to identify relationships among particular variables. Therefore, survey studies may vary in complexity, ranging from simple frequency counts to relational analyses (Cohen, Manion, & Morrison, 2007). Survey models are typically used to examine a broad subject area or population in order to measure or describe any generalized characteristic. According to Anderson (1990), as cited in Balci (2000), there are four types of research commonly conducted in educational settings: descriptive, causal, generalization, and fundamental or theoretical research. Considering these classifications, the present study can be categorized as a "generalization" type of research. Generalization studies aim to extend findings derived from causal research to new contexts or situations (Shadish, Cook, & Campbell, 2002).

Participants

The sample of the study consisted of 598 students enrolled in public primary schools located on the European side of Istanbul. The sample was determined using the convenience sampling method. Convenience sampling is based on selecting participants from units that are accessible and feasible for the researcher (Büyüköztürk, Çakmak, Akgün, Karadeniz, & Demirel, 2012; Etikan et al., 2016). This method provides practical advantages in terms of time, cost, and accessibility. While using this sampling method, efforts were made to reach participants from different grade levels and with diverse demographic characteristics in order to reduce potential sampling bias.

The convenience sampling method was preferred in this study because collecting data based on voluntary participation was considered more practical and feasible. The sample of the study comprised primary school students and their parents from five schools located in the districts of Küçükçekmece, Avclar, and Bakirköy. Demographic information regarding the participants is presented in Table 1.

Of the 598 participants included in the study, 55.7% were female and 44.3% were male. The majority of the participants (67.9%) were third-grade students, while 32.1% were fourth-grade students. Regarding the number of siblings, the largest proportion of participants (30.6%) had two siblings, followed by those with one sibling (29.1%). In terms of mothers' educational level, the highest proportion (34.3%) consisted of primary school graduates, followed by middle school graduates (27.3%) and high school graduates (24.1%). Similarly, regarding fathers' educational level, the largest proportion (32.6%) consisted of high school graduates, followed by primary school graduates (27.8%) and middle school graduates (26.3%). Overall, the educational levels of the participants' parents were predominantly concentrated at the primary, middle, and high school levels.

Table 1.
Demographic Characteristics of the Participants

Variable	Category	n	%
Gender	Female	333	55,7
	Male	265	44,3
Grade Level	3rd Grade	406	67,9
	4th Grade	192	32,1
	0	35	5,9
Number of Siblings	1	174	29,1
	2	183	30,6
	3	145	24,2
	4	36	6,0
	5	19	3,2
	6	6	1,0
Mother's Educational Background	Uneducated	19	3,2
	Primary School	205	34,3
	Middle School	163	27,3
	High School	144	24,1
	Associate Degrees	17	2,8
	Bachelor's Degrees	49	8,2
	Postgraduate Degrees	1	0,2
	Uneducated	9	1,5
Father's Educational Background	Primary School	166	27,8
	Middle School	157	26,3
	High School	195	32,6
	Associate Degrees	19	3,2
	Bachelor's Degrees	49	8,2
	Postgraduate Degrees	3	0,5

Data Collection Instruments

Four data collection instruments were employed to gather the data for this study: the Creative Thinking Scale, the Critical Thinking Scale, the Parenting Style Scale, and the Perceived Stress Scale. The creative and critical thinking scales were administered to primary school students, whereas the parenting style and perceived stress scales were administered to the mothers or fathers of the same students in 2025. Quantitative data were obtained through the administration of these instruments, and subsequent analyses were conducted based on the collected data. Detailed information regarding the data collection instruments is presented below.

Creative Thinking Scale

In this study, the Creative Thinking Scale developed by Deniz and Demir (2024), with established construct validity, was employed. The scale was developed to determine the creative thinking potential of students aged 9–15. The construct validity of the scale was examined through both exploratory factor analysis (EFA) and confirmatory factor analysis (CFA), and the model fit indices indicated a satisfactory fit. The test-retest reliability coefficient of .71 demonstrates that the scale maintains measurement stability over time and yields consistent results. Furthermore, the internal consistency coefficient of the scale was calculated

as .758 based on the data obtained within the scope of this study. In addition, the difference analysis conducted between the upper 25% and lower 25% groups yielded statistically significant results [$t(249.295) = -47.28, p < .001$]. These findings indicate that the scale produces valid and reliable measurements. The scale consists of 12 items, is unidimensional, and is structured as a 5-point Likert-type scale. Additionally, in order to determine students' creative thinking levels based on the data obtained from the measurement tool, the level ranges specified below were utilized.

1.00–1.80 "Very Low"

1.81–2.60 "Low"

2.61–3.40 "Moderate"

3.41–4.20 "High"

4.21–5.00 "Very High"

Critical Thinking Scale

Another data collection instrument used in this study was the Critical Thinking Scale, whose construct validity was established by Döner and Demir (2022) through a study conducted with students aged 9–14. The scale consists of 21 items structured on a 5-point Likert-type format and includes three factors focusing on skills such as analysis, evaluation, and interpretation. Within the scope of the construct validity study, different sample groups were employed independently at each stage of the analysis (EFA, CFA, and test-retest). Since a second-order confirmatory factor analysis (CFA) was performed, the scale demonstrates a structure suitable for use based on the total score. Based on the data obtained in the present study, the internal consistency coefficient of the scale was calculated as .832. The reliability coefficients for the sub-dimensions were found to be .781 for the disposition sub-dimension, .580 for the analysis sub-dimension, and .822 for the dialectical sub-dimension. Furthermore, analyses conducted between the upper and lower groups revealed a statistically significant difference between the mean scores of the groups [$t(264.157) = -43.21, p < .001$]. These findings indicate valid and reliable measurements. In addition, in order to determine students' critical thinking tendencies based on the data obtained from the measurement tool, the level ranges were as follows.

1.00–1.80 "Very Low"

1.81–2.60 "Low"

2.61–3.40 "Moderate"

3.41–4.20 "High"

4.21–5.00 "Very High"

Perceived Stress Scale

Another data collection instrument used in this study was the Perceived Stress Scale. Developed by Snoeren-Hoefnagels (2014), the scale consists of nine items with a single-dimensional structure. The original

scale was administered to 223 students (100 girls and 123 boys) aged between 8 and 11 years. Higher scores on the scale indicate higher levels of perceived stress among individuals. The scale was adapted into Turkish by Oral and Ersan (2017). As part of the adaptation study, construct validity analyses, criterion-related validity, and reliability analyses, including test-retest reliability and Cronbach's alpha internal consistency coefficient, were conducted. Based on these analyses, the scale was reported to be a valid and reliable measurement tool and considered appropriate for determining perceived stress levels in children aged 8–11 years. Within the scope of the present study, the internal consistency coefficient of the scale was calculated as .732. Additionally, analyses conducted between the upper and lower groups revealed a statistically significant difference between the mean scores of the groups [$t(224.966) = -46.078, p < .001$]. These findings indicate that the scale produces valid and reliable measurements. Furthermore, in order to determine students' stress levels based on the data obtained from the measurement instrument, the level ranges specified below were utilized.

- 1.00–1.75 "Low Stress Level"
1.76–2.50 "Lower-Moderate Stress Level"
2.51–3.25 "Upper-Moderate Stress Level"
3.26–4.00 "High Stress Level"

Multidimensional Parenting Styles Assessment Scale

In this study, the Multidimensional Parenting Styles Assessment Scale, adapted into Turkish by Karababa (2019), was used to determine parenting styles. In the adaptation study of the scale, exploratory factor analysis (EFA) and confirmatory factor analysis (CFA) were conducted for validity and reliability purposes. In addition, criterion-related validity, internal consistency, and test-retest reliability analyses were performed, and the related coefficients were found to

be at acceptable levels. The scale consists of 34 items and seven sub-dimensions: (1) Proactive Parenting (6 items), (2) Positive Reinforcement (4 items), (3) Intimate Relationship (3 items), (4) Supportive Approach (3 items), (5) Hostility (7 items), (6) Low Control (7 items), and (7) Physical Control (4 items). The scale can also be evaluated by calculating total scores for two dimensions: positive parenting (proactive parenting, positive reinforcement, warmth, and supportiveness) and negative parenting (hostility, low control, and physical control) (Karababa, 2019).

In the present study, the internal consistency values of the factors were examined separately. The internal consistency coefficients were examined separately for these two dimensions, yielding coefficients of .766 for comprehensive positive parenting and .711 for comprehensive negative parenting. These findings indicate that the internal consistency coefficients of both dimensions are above the acceptable threshold of .60 (Akbulut, 2010, Hair et al., 2013).

Data Analysis

Data analyses were conducted using IBM SPSS Statistics 25 and Jamovi. Initially, missing data and outlier analyses were performed. Descriptive statistics, including frequency, percentage, mean, and standard deviation values, were calculated. Independent samples t-test and one-way analysis of variance (ANOVA) were used for comparisons between groups. The level of statistical significance was set at .05. The normality of the data distribution was assessed by examining skewness and kurtosis values. Decisions regarding the use of parametric or non-parametric analyses were made based on the normality distribution of the independent variables. In this context, the following statistical values and analysis techniques were included in the study.

Table 2

Skewness-kurtosis values for measurement instruments and independent variables

Variable Category		Perceived Stress Scale		Creative Thinking Scale		Critical Thinking Scale	
		Skewness	Kurtosis	Skewness	Kurtosis	Skewness	Kurtosis
Gender	Female	,309	-,168	-,341	-,023	-,441	,389
	Male	,164	-,409	-,414	-,284	-,155	-,166
Grade Level	3rd Grade	,228	-,300	-,359	-,086	-,341	-,010
	4th Grade	,251	-,258	-,396	-,263	-,152	,085
	Uneducated	-,028	,576	,157	-,137	-,727	,528
Mother's Educational Background	Primary School	,162	-,302	-,448	-,186	-,338	,241
	Middle School	,346	-,406	-,135	-,233	-,293	-,299
	High School	,143	-,209	-,369	-,145	-,062	,117
	Associate Degree	1,402	2,846	-,257	6,563	-,1048	,829
	Bachelor's/Postgraduate Degree	,264	,089	-,013	-,810	-,023	-,1115
Father's Educational Background	Uneducated	-,085	,139	-,406	-,183	-,176	2,093
	Primary School	,230	-,332	-,349	-,341	-,265	,106
	Middle School	,246	-,169	-,237	-,142	-,542	,604
	High School	,138	-,410	-,435	-,034	-,236	-,261
	Associate Degree	,205	,277	-,497	,029	,109	-,1327
Overall Scale	Bachelor's/Postgraduate Degree	,595	,041	-,699	,968	-,688	,563
		,230	-,300	-,356	-,141	-,267	,043

Table 2 indicates that the data obtained from the measurement instruments demonstrated normal distribution within acceptable limits (-2 to $+2$) in terms of gender and grade level (George & Mallery, 2010). However, the skewness and kurtosis values for mothers' education level fell outside the acceptable range for the creative thinking and perceived stress scales. Therefore, non-parametric analyses were employed for the relevant measurements. It can also be observed that the skewness and kurtosis values for the overall scale scores were within acceptable limits. Additionally, Jamovi version 2.6.44 and the PROCESS Macro 5.0 statistical programs were used to examine the relationship between parenting styles and perceived stress among primary school students and to evaluate this relationship with the mediating roles of creative and critical thinking tendencies.

Within the scope of the study, the relationships among perceived stress, creative thinking, critical thinking, and parenting styles were examined using mediation models. First, the mediating role of critical thinking in the relationship between perceived stress and creative thinking was tested. Subsequently, the mediating role of creative thinking in the relationship between perceived stress and critical thinking was examined. In addition, the mediating effects of creative thinking and critical thinking on the relationship between parenting styles and perceived stress were evaluated separately. Finally, a serial mediation model including both creative thinking and critical thinking was tested.

Mediation analyses were conducted using PROCESS Macro, and PROCESS Model 4 was employed for the simple mediation analyses. The significance of indirect effects was evaluated using the bootstrap method based on 5000 resamples, and 95% confidence intervals were taken into account. Confidence intervals that did not include zero were considered indicative of a significant indirect effect. The findings were interpreted within the limitations of the cross-sectional design, and therefore no causal inferences were made.

Validity and Reliability of the Data

First, careful consideration was given to ensuring that the measurement instruments used in the study were appropriate for the sample and that the necessary construct validity studies had been conducted. In

addition, internal consistency coefficients and upper-lower group discrimination indices were examined to assess the reliability and validity of the measurement data. The results of these analyses were presented in the sections describing the measurement instruments. The obtained values indicated that the measurements produced valid and reliable results.

Ethical Approval of the Research

The study was conducted following ethical approval obtained from the İstanbul Kültür University Ethics Committee, in accordance with the decision dated April 21, 2025, and numbered 2025/17.

Findings

This section presents the findings obtained from the research data along with relevant interpretations. The findings are presented in accordance with the order of the research questions.

Research problem 1: What are the perceived levels of stress, creative thinking, and critical thinking among students?

The first objective of the study was to examine students' levels of perceived stress, creative thinking, and critical thinking. In this context, interpretations regarding students' levels were made based on arithmetic means and predetermined level ranges. The value intervals specified in the measurement instruments were used to determine these level ranges.

As shown in Table 3, students' perceived stress levels were observed to be "low-medium," while their creative and critical thinking levels were "high."

Research problem 2: What parenting styles do parents exhibit?

The second sub-objective of the study was to examine the parenting styles of the parents. In this context, the parenting style with the highest arithmetic mean score in the data collection instrument was identified as the dominant parenting style for each participant. The distribution of parenting styles is presented in Table 4.

Table 3.
Students' perceived stress, creative thinking, and critical thinking levels

	N	Min	Max	Mean	Standard deviation	Level
Students' perceived stress	598	1,00	4,00	2,2573	,602	Lower Middle
Creative thinking	598	1,75	5,00	3,6199	,598	High
Critical thinking	598	1,86	5,00	3,6937	,595	High

Table 4.

Parenting Styles

Parenting Style	Total Number of Participants	%
Positive Reinforcement	264	44,1
Warmth	200	33,4
Supportiveness	123	20,6
Proactive Parenting	11	1,8
Total	598	100

Table 4 indicates that the majority of parents (44.1%) demonstrated a positive reinforcement parenting style. No parents exhibited physical control, lax control, or hostile parenting styles. The analysis of the dataset also revealed that some parents demonstrated characteristics of more than one parenting style; in such cases, parents were classified according to their most dominant parenting style.

Research problem 3: Do students' perceived stress levels differ according to their families' parenting styles?

Within the scope of this sub-objective, the effect of family parenting styles on students' perceived stress levels was examined. Since the significance values obtained from the test of homogeneity of variances were greater than .05 (ranging between .254 and .306), the variances were considered homogeneous. Therefore, one-way analysis of variance (ANOVA) was conducted. The results of the analysis are presented below.

Table 5.

The effect of parenting styles on perceived stress ANOVA

	Sum of Squares	df	Mean Square	F	p
Between Groups	2,862	-3	,954	2,653	,048
Within Groups	213,640	594	,360		
Total	216,502	597			

Figure 1.

Mediation model illustrating the mediating role of creative thinking between parenting styles and perceived stress

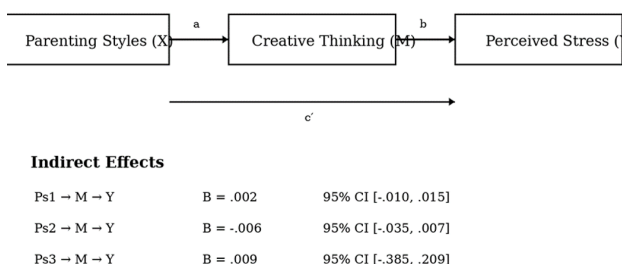


Figure 1 presents the mediation model examining the mediating role of creative thinking in the relationship between parenting styles and perceived stress.

Although one direct pathway was statistically significant, none of the indirect effects were significant, as the confidence intervals included zero.

Table 6.

The mediating role of creative thinking between parental style and perceived stress

Effect	B	se	t	p	LLCI	ULCI
X→M						
Ps1	,027	,056	,487	,626	-,083	,138
Ps2	-,073	,065	-1,120	,263	-,201	,055
Ps3	-,020	,184	-,109	,913	-,382	,342
R ²	.004	F .734		P=.532		
X→Y						
Ps1	-,055	,056	-,979	,328	-,165	,055
Ps2	-,110	,065	-1,673	,095	-,238	,019
Ps3	,367	,185	1,988	,047	,004	,729
R ²	.013	F=2,653		P=.048		
Indirect Path		B	BootSE	LLCI	ULCI	
Ps1 → M → Y		,002	,006	-,010	,015	
Ps2 → M → Y		-,006	,010	-,035	,007	
Ps3 → M → Y		,009	,182	-,385	,209	

Note: Ps --- Parenting style

When the table is examined, the findings of the mediation analysis conducted using the bootstrap method (5,000 samples) indicate that the indirect effects of variables X1, X2, and X3 on the dependent variable were not statistically significant. Since all confidence intervals included zero (X1: [-.010, .015]; X2: [-.035, .007]; X3: [-.385, .209]), it was concluded that the mediation effects were not statistically significant.

When the total effects were examined, only variable X3 demonstrated a statistically significant effect on the dependent variable (B = .367, p = .047). However, this effect lost its statistical significance in the direct effect model (p = .053). Furthermore, the non-significant indirect effects indicate that the mediation relationship was not supported.

Research problem 5: Does critical thinking tendency play a significant mediating role in the relationship between parenting styles and perceived stress?

This research question examined the mediating role of critical thinking tendency. The findings are presented below.

Figure 2.

Mediation model of critical thinking in the relationship between parenting styles and perceived stress

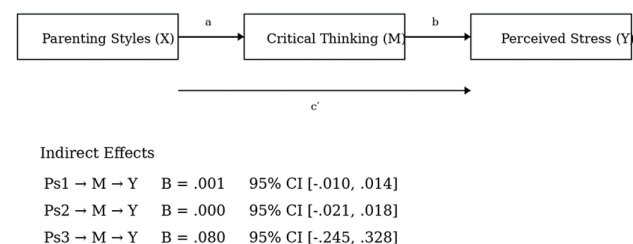


Figure 2 presents the mediation model examining the mediating role of critical thinking in the relationship between parenting styles and perceived stress. Since the confidence intervals for the indirect effects included zero, the mediation effect was not found to be statistically significant.

Table 7.

Mediating role of critical thinking between parental styles and perceived stress

Effect	B	se	t	p	LLCI	ULCI
X--M						
Ps1	,013	,056	,237	,812	-,096	,123
Ps2	-,072	,065	-1,105	,269	-,199	,056
Ps3	-,144	,183	-,789	,431	-,504	,215
R ² .004	F .751		P=,522			
Direct X--Y						
Ps1	-,055	,056	-,979	,328	-,165	,055
Ps2	-,110	,065	-1,673	,095	-,238	,019
Ps3	,367	,185	1,988	,047	,004	,729
R ² =.013	F=2,653		P=,048			
Indirect Path			B	BootSE	LLCI	ULCI
Ps1 → M → Y			,001	,005	-,010	,014
Ps2 → M → Y			,000	,009	-,021	,018
Ps3 → M → Y			,080	,154	-,245	,328

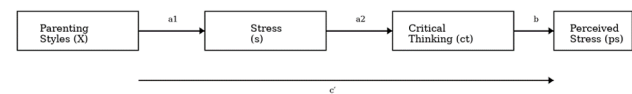
The results of the mediation analysis indicated that the effects of the independent variables on the mediating variable were not statistically significant ($p > .05$). Furthermore, the indirect effects obtained using the bootstrap method showed that all confidence intervals included zero (X1: [-.010, .014]; X2: [-.021, .018]; X3: [-.245, .328]). These findings indicate the absence of a mediating effect in the model. When the direct effects were examined, only variable X3 was found to have a statistically significant effect on the dependent variable ($B = .367$, $p = .047$). However, this effect was not observed to occur through the mediating variable. In conclusion, the findings do not support a mediating relationship in the examined model.

Research problem 6: Do creative and critical thinking tendencies together play a significant mediating role in the relationship between parenting styles and perceived stress?

Within the scope of this objective, creative thinking and critical thinking tendencies were considered together as mediating variables to examine their mediating relationships.

Figure 3.

Serial mediation model examining indirect pathways among variables



Indirect Effects

Ps1 → s → ct → ps	B = -.001	95% CI [-.015, .013]
Ps2 → s → ct → ps	B = .007	95% CI [-.006, .029]
Ps3 → s → ct → ps	B = .014	95% CI [-.027, .070]
Ps1 → s → Crt → ps	B = .001	95% CI [-.006, .011]
Ps2 → s → Crt → ps	B = -.002	95% CI [-.015, .007]
Ps3 → s → Crt → ps	B = .006	95% CI [-.011, .031]
Ps1 → s → ct → Crt → ps	B = .001	95% CI [-.008, .009]
Ps2 → s → ct → Crt → ps	B = -.004	95% CI [-.019, .004]
Ps3 → s → ct → Crt → ps	B = -.008	95% CI [-.044, .017]

Figure 3 presents the serial mediation model examining the relationships among parenting styles, stress, critical thinking, and perceived stress. The findings indicated that none of the indirect pathways were statistically significant, as all bootstrap confidence intervals included zero. The findings obtained in this regard are presented in Table 8:

Table 8.

Mediating roles of creative and critical thinking tendencies between parental styles and perceived stress

Indirect Path	B	BootSE	LLCI	ULCI
Ps1 → s → ct → ps	-,001	,006	-,015	,013
Ps2 → s → ct → ps	,007	,009	-,006	,029
Ps3 → s → ct → ps	,014	,023	-,027	,070
Ps1 → s → Crt → ps	,001	,004	-,006	,011
Ps2 → s → Crt → ps	-,002	,005	-,015	,007
Ps3 → s → Crt → ps	,006	,010	-,011	,031
Ps1 → s → ct → Crt → ps	,001	,004	-,008	,009
Ps2 → s → ct → Crt → ps	-,004	,006	-,019	,004
Ps3 → s → ct → Crt → ps	-,008	,015	-,044	,017

Note: Ps: Parenting style; s: style; ct: critical thinking; Crt: creative thinking; ps: perceived stress

When the bootstrap results of the serial mediation model were examined, it was observed that the confidence intervals for all indirect effect pathways included zero. This finding indicates that neither the specific indirect paths ($s \rightarrow ct \rightarrow ps$; $s \rightarrow Crt \rightarrow ps$) nor the full serial pathway ($s \rightarrow ct \rightarrow Crt \rightarrow ps$) were statistically significant. Therefore, it was concluded that the mediating variables in the model did not significantly mediate the relationship between the independent and dependent variables.

Discussion

This study focuses on the mediating roles of creative and critical thinking dispositions in the relationship between parenting styles and perceived stress. The findings obtained within the scope of the first objective of the study indicate that primary school students' perceived stress levels are at a "low-to-moderate" level. Based on this finding, it can be suggested that children in primary school experience

a certain degree of stress in their daily lives; however, this stress does not reach high or clinical levels. This result is consistent with various studies in the literature. For instance, Davis and Turner-Cobb (2023), in a study conducted with children aged 7–11, reported that the mean perceived stress score of the sample was close to the theoretical midpoint of the scale. Similarly, Eroen (2022), in a study involving 852 children aged 9–11, found that primary school students had low levels of perceived stress. Likewise, Lovenjak and Peklaj (2016) demonstrated that primary school students experienced school-related stress at low levels. One possible explanation for the relatively low perceived stress levels observed in the present study may be that, due to their age, children tend to perceive stress in a more superficial and short-term manner. According to the stress and coping theory proposed by Lazarus and Folkman (1984), individuals interpret environmental demands through cognitive appraisal processes and evaluate the adequacy of personal and environmental resources available to cope with these demands. Accordingly, how individuals perceive and evaluate such demands plays a crucial role in determining their stress levels (as cited in Folkman, 2013). Since children's cognitive capacities during the primary school period may not yet be fully developed. Additionally, according to the transactional theory of stress proposed by Richard Lazarus and Susan Folkman, stress is shaped by individuals' cognitive appraisal of a situation. It may therefore be argued that children at these ages experience relatively low levels of stress, since expectations arising from school and family environments are generally not perceived as threatening.

Another finding of the study indicates that students exhibit high levels of critical thinking disposition. A substantial body of research in the literature reports findings consistent with the results of the present study (Bakır & Eğmir, 2022; Demir, 2006; Oğuz & Aybek, 2024). However, there are also studies suggesting that primary school students' critical thinking dispositions are at a moderate level (Akar & Kara, 2016; Hocioğlu, 2020; Kiran, 2019; Ocak & Kalender, 2017). For instance, Yetgin and Katrancı (2020) reported that fourth-grade primary school students' critical thinking skills were at a moderate level, approaching the lower threshold. In line with the findings on critical thinking, the present study also revealed that students' creative thinking dispositions are at a high level. Similarly, Deniz and Demir (2024), in their validity and reliability study of the Creative Thinking Disposition Scale developed for students aged 9–15 in Türkiye, found that the mean scores of this age group were very close to those obtained in the present study. Işık and Uysal (2015), in a study comparing grade levels, demonstrated that third- and fourth-grade students scored significantly higher in flexibility, originality, and overall creative thinking than first- and second-grade students. On

the other hand, some studies indicate that students' creative thinking dispositions tend to decline as grade level increases (Atkıncı, 2001). The similarity observed between students' levels of creative and critical thinking dispositions in the present study may be explained by the conceptual overlap between these two types of thinking. While Facione (1990) defines critical thinking in terms of cognitive skills such as analysis, evaluation, inference, and explanation, Torrance (1974) and Guilford (1967) conceptualize creative thinking based on divergent thinking capacity. These two modes of thinking can be considered complementary, with development in one likely to support and enhance the other. Supporting this view, Yılmaz (2025), in a quasi-experimental study conducted with 80 middle school students in Kastamonu, found that design-based research activities significantly increased both creative and critical thinking dispositions compared to the control group. It was observed that, through design processes, students simultaneously employed creative idea generation and critical evaluation skills while working on real-life problem situations. In addition, the fact that the measurement instruments used in the study were based on self-report may also have contributed to the relatively high levels of these thinking skills reported by the students. The literature likewise suggests that students may evaluate their own thinking skills more positively than their actual level of competence, and that this tendency may stem from self-assessment biases and a lack of metacognitive awareness (Dunning et al., 2003).

According to the findings of the study, nearly all parents exhibited parenting styles that can be characterized as positive, namely "positive reinforcement," "warmth" and "supportiveness" The proactive parenting style was observed at a rate of only 1.8%, and no parents were identified as exhibiting "hostility," "physical control," or "low control" styles. This finding appears to differ, at least partially, from those reported in the literature. A review of studies on parental attitudes indicates that, alongside democratic parenting styles, negative parenting styles such as authoritarian, overprotective, or neglectful approaches are also observed at certain rates (Özdemir & Yalçın, 2021; Teke et al., 2020). This discrepancy may be attributed to structural differences in measurement instruments, the homogeneity of the sample, or participants' tendencies toward social desirability. In self-report-based measurements, social desirability bias can be a determining factor. Particularly in domains subject to social judgment, such as parenting practices, individuals may tend to provide responses aligned with normative expectations rather than their actual behaviors (Paulhus, 2002). The finding that proactive parenting was observed at a rate of only 1.8% is noteworthy. Proactive parenting involves supporting children in preparing for potential future challenges, developing strategies for managing unexpected

situations, and engaging in advance discussions with children about expectations and requirements (Parent & Forehand, 2017). Based on this result, proactive parenting, which represents a future-oriented and structured approach to parenting, has not yet gained widespread adoption among parents.

The findings of the study indicate that parenting style has a statistically significant but small effect on students' perceived stress levels. A review of the literature reveals that studies directly examining the relationship between parenting styles and perceived stress are limited. Some studies indicate that authoritarian and negative parenting styles lead individuals to perceive situations as more threatening, thereby contributing to higher levels of stress and avoidance responses (Repetti et al., 2002, as cited in Özdemir & Yalçın, 2021). A review of the literature also indicates that the effects of parenting on anxiety and stress become most evident when negative parenting styles are compared with positive parenting styles (Wolfradt et al., 2003; Zeevi-Cousin & Lavenda, 2023). In the present study, the absence of a significant relationship at the level of parenting style subdimensions implies that the effect of parenting may be indirect and limited rather than direct and strong. On the other hand, the absence of a significant relationship may be attributable to the fact that the sample largely consisted of individuals exhibiting positive parenting styles. This situation may have restricted the variance in the parenting variable. In the literature, this phenomenon is referred to as "range restriction," which may reduce correlation coefficients and cause relationships between variables to appear weak or non-significant (Quinn McNemar, 1969, as cited in Baron & Kenny, 1986). Therefore, the present findings should not necessarily be interpreted as evidence for the absence of such relationships, but rather as a possible reflection of a methodological limitation.

In the present study, the results of the mediation analysis revealed that creative thinking disposition does not play a significant mediating role in the relationship between parenting styles and perceived stress. These findings also indicate that parenting style does not have a significant effect on creative thinking. It can be argued that this result differs from those reported in the literature. Studies demonstrating that parenting styles or parental attitudes are significant predictors of creative thinking skills (Deng et al., 2023; Dong et al., 2022; Fan & Zhang, 2014; Kaya et al., 2012; Wang et al., 2023; Zhao & Yang, 2021) generally involve samples consisting of middle school, high school, or university students. In contrast, studies including primary school students remain limited in number. In the meta-analysis conducted by Fan and colleagues (2024), which included students ranging from preschool to university level, a small but statistically significant positive correlation was found between

parental involvement and student creativity. On the other hand, Li and Li (2025), in a large-scale study involving 10- and 15-year-old Chinese students, found that the quality of parent-child relationships did not exert a significant direct effect on students' self-reported creativity, suggesting that parental relational support may shape creative development primarily through indirect psychological pathways rather than through direct influence. In the present study, however, students were found to exhibit high levels of creative thinking, and these levels appeared to emerge independently of parenting style. This finding suggests that creative thinking may be understood as a dynamic and socially grounded competence shaped not only by the family environment but also by the quality of learning experiences. Consistent with this interpretation, the literature provides evidence that school environments and pedagogical approaches play a decisive role in the development of creative thinking among primary school children (Fauzi & Wuryandani, 2025; Işık & Uysal, 2015; Xu et al., 2026). In a comprehensive systematic review and bibliometric analysis of 204 Scopus-indexed articles published between 2014 and 2024, Fauzi and Wuryandani (2025) demonstrated that creative thinking in primary education is fostered through constructivist, inquiry-based STEM, arts-integrated, and digital-hybrid pedagogical approaches.

It is also noteworthy that creative thinking did not have a significant effect on perceived stress. Russ (2014) suggest that creative thinking can facilitate individuals' coping processes and enhance emotional flexibility. However, this protective effect is reported to be more evident particularly among adolescents or adults whose cognitive development has largely matured. Coping in primary school-aged children is considered a developmental process that involves conscious and voluntary efforts to regulate emotion, cognition, and behavior (Compas et al., 2001). Within this framework, since these regulatory capacities develop gradually throughout childhood, it can be argued that children in the primary school period may not yet have fully developed the cognitive self-regulation skills necessary to use their creative thinking dispositions as functional and deliberate coping strategies. Accordingly, it can be suggested that the link between creative thinking disposition and stress management has not yet fully matured in this age group.

The results of the mediation analysis also revealed that critical thinking disposition does not play a significant mediating role in the relationship between parenting styles and perceived stress. This finding indicates that there is no significant association between parenting styles and critical thinking, and therefore, the fundamental condition required for a mediation effect is not met. This result appears to differ from findings

reported in the literature (Deng et al., 2023; Wang et al., 2020; Yurt & Kara, 2025). Deng et al. (2023), in a study conducted with primary school students, found that parenting styles were significantly associated with critical thinking dispositions, and longitudinal analyses showed that parenting styles predicted these dispositions over time. In the present study, although students were found to have high levels of critical thinking disposition, the lack of a significant effect of parenting styles suggests that school-based practices may play a more influential role in the development of this disposition and related skills. Consistent with this interpretation, the literature indicates that practices such as reading, research-oriented and inquiry-based instructional methods, as well as school climate, are effective in fostering these skills (Dimaç & Dimaç, 2020; Gómez & Suárez, 2020; Karataş & Özcan, 2015; Oğuz & Aybek, 2024; Türkdoğan & Özgenel, 2021; Yılmaz, 2025). Yıldırım Döner and Demir (2022) also emphasize that both family environment and school- and teacher-related factors should be considered among the determinants of critical thinking disposition. They suggest that parental attitudes at home may influence the extent to which individuals can express themselves, while teachers and schools contribute to the development of this disposition by creating classroom environments that support critical thinking.

The finding of the present study that critical thinking does not have a reducing effect on children's perceived stress appears to be consistent with the results reported by Arıbaş and Oğuzhan (2024). In their study conducted with university students, Arıbaş and Oğuzhan (2024) found a negative but weak relationship between critical thinking and perceived stress; however, regression analyses indicated that critical thinking did not significantly predict perceived stress levels. These findings suggest that the relationship between critical thinking and perceived stress remains weak and limited. On the other hand, Arrivillaga et al. (2022) reported a negative relationship between critical thinking and perceived stress. In the present study, the absence of a significant relationship between children's critical thinking dispositions and perceived stress may be attributed to developmental characteristics specific to this age group. According to Halpern (2014), critical thinking encompasses higher-order cognitive processes such as tolerance for ambiguity, flexibility, self-regulation, and complex problem-solving. These processes can be considered to still be in the process of maturation during the primary school period. Lazarus and Folkman (1984), in their stress and coping model, identified cognitive reappraisal as one of the most effective strategies for managing stressful situations. Critical thinking skills may also be associated with cognitive reappraisal processes (as cited in Folkman, 2013). However, in coping with perceived stress, children tend to rely more on externally oriented strategies rather than

cognitive evaluations. Yıldırım Döner and Demir (2022) also include age level among the factors influencing critical thinking disposition and, drawing on Piaget (2019), emphasize that perceptual activities and related thinking skills develop both quantitatively and qualitatively with age. Coping strategies also change developmentally. Goodman and Thompson (2010) indicate that cognitively based coping strategies, such as cognitive reappraisal, begin to be used effectively during middle childhood; prior to this period, children tend to rely primarily on behavioral strategies such as seeking social support, avoidance, or distraction when dealing with stressful situations (as cited in Compas et al., 2020). In addition, it should be taken into consideration that the scales used in the study measured dispositions toward creative and critical thinking rather than directly assessing performance-based indicators of thinking. It may therefore be argued that, in order to function as a mediating factor in reducing stress in a meaningful way, such dispositions need to be translated into actual behaviors and active coping strategies.

The serial mediation analysis conducted in the study revealed that creative and critical thinking dispositions, even when considered jointly, do not play a significant mediating role in the relationship between parenting styles and perceived stress. This finding is consistent with the results obtained in the previous sub-objectives and indicates that the proposed mediation mechanism does not operate even when both thinking dispositions are included simultaneously. Based on this result, it can be suggested that creative and critical thinking dispositions are shaped by multiple determinants. In other words, these thinking dispositions cannot be considered unidimensional constructs that can be explained solely by parenting characteristics, nor are they sufficient on their own to transmit the effects of parenting styles to stress outcomes.

Implications

The findings of the study revealed that although students exhibit high levels of creative and critical thinking dispositions, these dispositions do not function effectively in stress management. This may indicate that such skills are addressed in school settings primarily within the scope of academic achievement. Designing instructional activities aimed at developing these skills in a way that also encompasses emotional regulation, coping strategy development, and self-awareness may facilitate the integration of thinking dispositions into everyday life. From the perspective of parenting style, it is noteworthy that although the majority of parents in the sample demonstrated positive parenting styles, proactive parenting was observed at a very limited rate. In this regard, it is recommended that family education programs

organized in schools introduce this approach through concrete examples, and that teachers and school counselors share strategies during parent meetings that support preparing children for potential future challenges. On the other hand, considering the finding that the effect of parenting style on student stress is limited, school-based stress reduction interventions may yield more effective outcomes if implemented from an integrated perspective that extends beyond parenting styles to include school climate, teacher-student relationships, and peer interactions.

Since the present study is limited to the primary school level, testing similar models with middle and high school students is expected to allow for a comparative examination of the relationships among parenting styles, thinking dispositions, and perceived stress from a developmental perspective. The findings indicated that the effect of parenting style on stress is limited and indirect; therefore, future research may explore the potential mediating or moderating roles of variables such as emotional regulation, self-efficacy, and peer relationships. However, the cross-sectional design employed in this study restricts the ability to establish causal relationships. The use of longitudinal or experimental designs may enable a more robust examination of the relationships among variables within a causal framework.

Conclusion

This study examined the mediating roles of creative and critical thinking dispositions in the relationship between parenting styles and perceived stress among primary school students. The findings revealed that students' perceived stress levels were at a low-to-moderate level, while their creative and critical thinking dispositions were at a high level. Parenting style was found to have a statistically significant but small effect on perceived stress. Mediation analyses indicated that there was no significant mediating effect for creative thinking, critical thinking, or the serial mediation model in which both variables were considered together. Overall, the study provides important evidence emphasizing that the factors influencing stress in primary school children should be addressed within a multidimensional and developmental framework.

AI Disclosure Statement

The authors confirm that no artificial intelligence tools were used at any stage of the study.

References

- Akar, C. & Kara, M. (2016). İlkokul 4. sınıf öğrencilerinin eleştirel düşünme becerilerinin bazı değişkenlere göre değerlendirilmesi [An evaluation of fourth-grade primary school students' critical thinking skills according to various variables]. *Uluslararası Türkçe Edebiyat Kültür Eğitim (Teke) Dergisi*, 5(3), 1339-1355. <https://dergipark.org.tr/tr/pub/teke/article/263928>
- Akbulut, Y. (2010). Sosyal bilimlerde SPSS uygulamaları: Sık kullanılan istatistiksel analizler ve açıklamalı SPSS çözümleri [SPSS applications in the social sciences: frequently used statistical analyses and annotated SPSS solutions]. İdeal Kültür & Yayıncılık.
- Anderson, L. W., & Krathwohl, D. R. (2001). *A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives (Complete ed.)*. Addison Wesley Longman.
- Arıbaş, A. N., & Oğuzhan, Y. S. (2024). Eleştirel düşünmenin problem çözme ve algılanan stres düzeyine etkisi [The effect of critical thinking on problem-solving and perceived stress levels]. *Manisa Celal Bayar Üniversitesi Sosyal Bilimler Dergisi*, 22(1), 272-289. <https://doi.org/10.18026/cbayarsos.1423629>
- Arrivillaga, C., Rey, L., & Extremera, N. (2022). The chain-mediating role of perceived stress and problematic smartphone use in the link between critical thinking and academic engagement among Spanish adolescents. *Cyberpsychology: Journal of Psychosocial Research on Cyberspace*, 16(4). <https://doi.org/10.5817/CP2022-4-7>
- Atkinci, H. (2001). *İlköğretim birinci kademe eğitim programlarının yaratıcı düşünmenin gelişmesine etkileri* [The effects of primary school curricula on the development of creative thinking, Unpublished master's thesis]. Çanakkale Onsekiz Mart University, Türkiye.
- Bakır, T., & Eğmir, E. (2022). Examination of the relationship between secondary school students' critical thinking dispositions and metacognitive awareness. *E-International Journal of Educational Research*, 13(5), 21-40. <https://doi.org/10.19160/e-ijer.1105032>
- Balcı, A. (2000). *Sosyal bilimlerde araştırma: Yöntem, teknik ve ilkeler* [Research in social sciences: Methods, techniques, and principles] (11th ed.). Pegem Akademi

- Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory*. Prentice-Hall.
- Baron, R. M., & Kenny, D. A. (1986). The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology*, 51(6), 1173–1182. <https://doi.org/10.1037/0022-3514.51.6.1173>
- Bronfenbrenner, U. (1979). *The ecology of human development: Experiments by nature and design*. Harvard University Press.
- Cohen, L., Manion, L., & Morrison, K. (2007). *Research methods in education* (6th ed.). Routledge
- Compas, B. E., Connor-Smith, J. K., Saltzman, H., Thomsen, A. H., & Wadsworth, M. E. (2001). Coping with stress during childhood and adolescence: Problems, progress, and potential in theory and research. *Psychological Bulletin*, 127(1), 87–127. <https://doi.org/10.1037/0033-2909.127.1.87>
- Crilly, N. (2025). Critical thinking, creative thinking, systems thinking and many more: A comparative bibliometric analysis of prevalence and distribution. *Thinking Skills and Creativity*, 56, 102014. <https://doi.org/10.1016/j.tsc.2025.102014>
- Davis, C., & Turner-Cobb, J. M. (2023). The Perceived Stress Scale for Kids (PeSSKi): Initial development of a brief measure for children aged 7–11 years. *Stress and Health*, 39(1), 125–136. <https://doi.org/10.1002/smi.3174>
- Demir, M. K. (2006). İlköğretim dördüncü ve beşinci sınıf öğrencilerinin sosyal bilgiler derslerinde eleştirel düşünme düzeylerinin çeşitli değişkenler açısından incelenmesi [Examination of fourth- and fifth-grade primary school students' critical thinking levels in social studies courses in terms of various variables]. *Gazi Üniversitesi Gazi Eğitim Fakültesi Dergisi*, 26(3), 155–169. <https://dergipark.org.tr/tr/pub/gefad/issue/6752/90797>
- Deniz, H. & Demir, S. (2024). Yaratıcı düşünme eğilimi ölçeği: Geçerlik ve güvenilirlik çalışması [Creative thinking disposition scale: Validity and reliability study]. *Gazi Üniversitesi Gazi Eğitim Fakültesi Dergisi*, 44(1), 703–735. <https://doi.org/10.17152/gefad.1347663>
- Deng, M., Ma, J., Lv, X., & Ren, X. (2023). Academic performance and parenting styles differentially predict critical thinking skills and dispositions among primary students: Cross-sectional and cross-lagged evidence. *Thinking Skills and Creativity*, 50, 101384. <https://doi.org/10.1016/j.tsc.2023.101384>
- Dilmaç, S., & Dilmaç, O. (2020). Görsel sanatlar dersinde yaratıcı drama yönteminin öğrencilerin eleştirel düşünme becerilerine etkisi [The effect of the creative drama method in visual arts courses on students' critical thinking skills]. *Yaratıcı Drama Dergisi*, 15(1), 85–100. <https://izlik.org/JA78WA35PY>
- Dong, Y., Lin, J., Li, H., Cheng, L., Niu, W., & Tong, Z. (2022). How parenting styles affect children's creativity: Through the lens of self. *Thinking Skills and Creativity*, 45, 101045. <https://doi.org/10.1016/j.tsc.2022.101045>
- Dunning, D., Johnson, K., Ehrlinger, J., & Kruger, J. (2003). Why people fail to recognize their own incompetence. *Current Directions in Psychological Science*, 12(3), 83–87. <https://doi.org/10.1111/1467-8721.01235>
- Ekşi, H., Çini, A., & Sevim, E. (2020). Çocuklarda okul temelli yalnızlık ve yaşam doyumu arasındaki ilişkide algılanan stresin aracı rolü [The mediating role of perceived stress in the relationship between school-based loneliness and life satisfaction in children]. *Kastamonu Education Journal*, 28(3), 1460–1470. <https://doi.org/10.24106/kefdergi.4128>
- Ercen, R. Ç. (2022). 9-11 yaş arası çocuklarda ekran kullanım düzeyi ile ilişkili faktörler; algılanan stres, yaşam doyumu, otomatik düşünceler [Factors associated with screen use levels in children aged 9–11 years: Perceived stress, life satisfaction, and automatic thoughts] [Unpublished doctoral dissertation] [Unpublished doctoral dissertation], İstanbul Arel University, Türkiye.
- Etikan, I., Musa, S. A., & Alkassim, R. S. (2016). Comparison of convenience sampling and purposive sampling. *American Journal of Theoretical and Applied Statistics*, 5(1), 1–4. <https://doi.org/10.11648/j.ajtas.2016050111>
- Facione, P. A. (1990). *Critical thinking: A statement of expert consensus for purposes of educational assessment and instruction*. American Philosophical Association. <https://files.eric.ed.gov/fulltext/ED315423.pdf>
- Fan, J., & Zhang, L.-F. (2014). The role of perceived parenting styles in thinking styles. *Learning and Individual Differences*, 32, 204–211. <https://doi.org/10.1016/j.lindif.2014.03.004>
- Fan, H., Feng, Y., & Zhang, Y. (2024). Parental involvement and student creativity: A three-level meta-analysis. *Frontiers in Psychology*, 15, 1407279. <https://doi.org/10.3389/fpsyg.2024.1407279>

- Fauzi, W. N. A., Wuryandani, W., & Supartinah. (2025). Creative thinking in global primary education: Pedagogical innovations and learning outcomes through an integrated bibliometric and systematic review. *Social Sciences & Humanities Open*, 12, 102216. <https://doi.org/10.1016/j.ssaho.2025.102216>
- Folkman, S. (2013). Stress: Appraisal and coping. M. D. Gellman & J. R. Turner (Eds.), *Encyclopedia of Behavioral Medicine* (s.1913-1915). Springer. https://doi.org/10.1007/978-1-4419-1005-9_215
- Fonseca, A., Moreira, H., & Canavarro, M. C. (2020). Uncovering the links between parenting stress and parenting styles: The role of psychological flexibility within parenting and global psychological flexibility. *Journal of Contextual Behavioral Science*, 18, 59–67. <https://doi.org/10.1016/j.jobs.2020.08.004>
- George, D., & Mallery, P. (2010). *SPSS for Windows step by step: A simple guide and reference (10th ed.)*. Pearson.
- Gómez, R. L., & Suárez, A. M. (2020). Do inquiry-based teaching and school climate influence science achievement and critical thinking? Evidence from PISA 2015. *International Journal of STEM Education*, 7(1), 43. <https://doi.org/10.1186/s40594-020-00240-5>
- Grolnick, W. S., & Pomerantz, E. M. (2009). Issues and challenges in studying parental control: Toward a new conceptualization. *Child Development Perspectives*, 3(3), 165–170.
- Guilford, J. P. (1967). *The nature of human intelligence*. McGraw-Hill.
- Halimah, S., Sit, M., & Robingatin. (2024). Parents' involvement in their children's primary education: A phenomenological study of Indonesian mothers' experiences. *Cogent Education*, 11(1), 2420421. <https://doi.org/10.1080/2331186X.2024.2420421>
- Halpern, D. F. (2014). *Thought and knowledge: An introduction to critical thinking* (5. baskı). Psychology Press.
- Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2013). *Multivariate data analysis* (7th ed.). Pearson.
- Hocaoğlu, M. E. (2020). *İlkokul öğrencilerinin argüman haritaları kullanımlarının eleştirel düşünme becerilerine etkisi [The effect of primary school students' use of argument maps on their critical thinking skills]* [Unpublished master's thesis], Gaziantep University, Türkiye.
- Işık, A. D., Uysal, H., Akosmanoğlu, E. & Bilir, A. (2015). *İlkokul eğitiminin yaratıcı düşünme becerileri üzerindeki etkisi [The effect of primary school education on creative thinking skills]*. *INESJOURNAL – Uluslararası Eğitim Bilimleri Dergisi*, 2(4), 358–367. <https://doi.org/10.16991/INESJOURNAL.98>
- Karababa, A. (2019). Çok boyutlu ebeveynlik stillerini değerlendirme ölçeğinin (ÇESDÖ) (ebeveyn formu) Türkçeye uyarlanması: geçerlik ve güvenirlik çalışması [Adaptation of the Multidimensional Parenting Styles Assessment Scale (MPSAS) (Parent Form) into Turkish: A Validity and Reliability Study]. *Pamukkale Üniversitesi Eğitim Fakültesi Dergisi*, 46, 108-131. <https://doi.org/10.9779/pauefd.439949>
- Karasar, N. (2020). *Bilimsel araştırma yöntemi: Kavramlar, ilkeler, teknikler [Scientific research methods: Concepts, principles, and techniques]* (35th ed). Nobel Akademik Yayıncılık.
- Karataş, S., & Özcan, S. (2015). İşbirlikli öğrenme ortamındaki yaratıcı etkinliklerin öğrencilerin yaratıcı ve eleştirel düşünceleri ile akademik başarıları üzerine etkisi [The effects of creative activities in a cooperative learning environment on students' creative and critical thinking and academic achievement]. *Eğitim Teknolojisi Kuram ve Uygulama*, 5(2), 1–21. <https://doi.org/10.17943/etku.56500>
- Karçı, A., & Çelik, S. B. (2024). Ergenlerde algılanan stres, akademik öz yeterlik ve psikolojik sağlık arasındaki ilişkinin incelenmesi [Examination of the relationship between perceived stress, academic self-efficacy, and psychological resilience in adolescents]. *Türk Eğitim Bilimleri Dergisi*, 22(1), 186–214. <https://doi.org/10.37217/tebd.1324851>
- Kaya, A., Bozaslan, H., & Genç, G. (2012). Üniversite öğrencilerinin anne-baba tutumlarının problem çözme becerilerine, sosyal kaygı düzeylerine ve akademik başarılarına etkisi [The effect of university students' parental attitudes on their problem-solving skills, social anxiety levels, and academic achievement]. *Dicle Üniversitesi Ziya Gökalp Eğitim Fakültesi Dergisi*, 18, 208–225. <https://izlik.org/JA35WR77RK>
- Kıran, M. S. (2019). *İlkokul 4. Sınıf Öğrencilerinin Eleştirel Düşünme Beceri Düzeyleri ile Okuduğunu Anlama Başarı Durumları Arasındaki İlişki [The relationship between fourth-grade primary school students' critical thinking skill levels and their reading comprehension achievement]* [Unpublished master's thesis]. İnönü University, Türkiye.

- Li, Z., & Li, Q. (2025). Unraveling the impact of parent-child relationships on adolescent creativity: The mediating effects of persistence and resilience. *Current Psychology*, 44, 14697–14712. <https://doi.org/10.1007/s12144-025-08183-2>
- Liu, D., Lin, J., Zhang, J., & Luo, X. (2025). Relationships among self-study ability, critical thinking ability, cooperative ability, and problem-solving ability in Chinese undergraduate nursing students: An analysis of a longitudinal cohort via cross-lagged models. *Frontiers in Psychology*, 16, 1606156. <https://doi.org/10.3389/fpsyg.2025.1606156>
- Lovenjak, I., & Peklaj, C. (2016). Stress and perception of school satisfaction on a sample of Slovene primary school students. *Psihologijske Teme*, 25(3), 357–379.
- McLeod, B. D., Wood, J. J., & Weisz, J. R. (2007). Examining the association between parenting and childhood anxiety: A meta-analysis. *Clinical Psychology Review*, 27(2), 155–172. <https://doi.org/10.1016/j.cpr.2006.09.002>
- Moran, K. M., Turiano, N. A., & Gentzler, A. L. (2018). Parental warmth during childhood predicts coping and well-being in adulthood. *Journal of Family Psychology*, 32(5), 610–621. <https://doi.org/10.1037/fam0000401>
- Ocak, İ. & Kalender, K. MD (2017). 6. sınıf öğrencilerinin eleştirel düşünme becerilerinin çeşitli değişkenler açısından incelenmesi (Kütahya ili örneği) [Examination of sixth-grade students' critical thinking skills in terms of various variables: The case of Kütahya province]. *Kastamonu Eğitim Dergisi*, 25(4), 1587-1600.
- Oğuz, O. & Aybek, B. (2024). İlkokul öğrencilerinin eleştirel düşünme eğilimlerinin çeşitli değişkenler açısından incelenmesi [Examination of primary school students' critical thinking dispositions in terms of various variables]. *Siirt Journal of Education*, 4(2), 1-18. <https://doi.org/10.58667/sedder.1543114>
- Oral, T., & Ersan, C. (2017). Çocuklarda (8-11 yaş) algılanan stres ölçeği'nin Türkçe'ye uyarlama çalışması [Turkish adaptation study of the perceived stress scale for children (8–11 years)]. *Selçuk Üniversitesi Edebiyat Fakültesi Dergisi*, 37, 419-428. <https://doi.org/10.21497/sefad.328629>
- Özdemir, E. & Yalçın, H. (2021). Anne-baba tutumlarının ergenlerdeki depresyon, anksiyete ve stres düzeylerine etkisi [The effect of parental attitudes on adolescents' levels of depression, anxiety, and stress]. *Karatay Sosyal Araştırmalar Dergisi*, (7), 229–248.
- Parent, J., & Forehand, R. (2017). The multidimensional assessment of parenting scale (MAPS): Development and psychometric properties. *Journal of Child and Family Studies*, 26(1), 2136–2151. [10.1007/s10826-017-0741-5](https://doi.org/10.1007/s10826-017-0741-5)
- Paulhus, D. L. (2002). Socially desirable responding: The evolution of a construct. H. I. Braun, D. N. Jackson & D. E. Wiley (Eds.), *The role of constructs in psychological and educational measurement* (ss. 49–69). Lawrence Erlbaum Associates.
- Qian, G., Li, B., Xu, L., Ai, S., Li, X., Lei, X., & Dou, G. (2024). Parenting style and young children's executive function mediate the relationship between parenting stress and parenting quality in two-child families. *Scientific Reports*, 14(1), 8503. <https://doi.org/10.1038/s41598-024-59225-x>
- Quan, Y. (2025). *The impact of parenting styles and family factors on children's creativity: A critical review*. In *Proceedings of the 2nd International Conference on Applied Psychology and Marketing Management (APMM), Volume 1* (pp. 677–681). Singapore. <https://doi.org/10.5220/0014150700004942>
- Runco, M. A., & Acar, S. (2012). Divergent thinking as an indicator of creative potential. *Creativity Research Journal*, 24(1), 66–75. <https://doi.org/10.1080/10400419.2012.652929>
- Russ, S. W. (2014). *Pretend play in childhood: Foundation of adult creativity*. American Psychological Association.
- Ryan, R. M., & Deci, E. L. (2017). *Self-determination theory: Basic psychological needs in motivation, development, and wellness*. Guilford Press. <https://doi.org/10.1521/978.14625/28806>
- Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Houghton Mifflin.
- Soenens, B., & Vansteenkiste, M. (2010). A theoretical upgrade of the concept of parental psychological control: Proposing new insights on the basis of self-determination theory. *Developmental Review*, 30(1), 74–99. <https://doi.org/10.1016/j.dr.2009.11.001>

- Sümer, N., Gündoğdu Aktürk, E., & Helvacı, E. (2010). Anne-baba tutum ve davranışlarının psikolojik etkileri: Türkiye’de yapılan çalışmalara toplu bakış [Psychological effects of parental attitudes and behaviors: a comprehensive review of studies conducted in Türkiye]. *Türk Psikoloji Yazıları*, 13(25), 42–59.
- Tabachnick, B. G., & Fidell, L. S. (2019). *Using multivariate statistics* (7th ed.). Pearson.
- Teke, E., Yılmaz, E. & Sürücü, A. (2020). İlkokul öğrencilerinin algıladıkları anne baba tutumları ile sosyal kaygı düzeyleri arasındaki ilişkinin incelenmesi [Examination of the relationship between primary school students’ perceived parental attitudes and their levels of social anxiety]. *The Journal of International Lingual, Social and Educational Sciences*, 6(1), 24–38. <https://doi.org/10.34137/jilses.647511>
- Torrance, E. P. (1974). *Torrance tests of creative thinking: Technical-norms manual*. Scholastic Testing Service.
- Türkdoğan, M., & Özgenel, M. (2021). The relationship between teachers creative thinking dispositions and school climate. *E-International Journal of Educational Research*, 12(1), 190–213. <https://doi.org/10.19160/e-ijer.815200>
- Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Harvard University Press.
- Wang, Q. (2023). The effect of parenting practices on creativity: Mediating role of psychological resilience. *Psychology Research and Behavior Management*, 16, 4501–4514. <https://doi.org/10.2147/PRBM.S436370>
- Wang, Y., Nakamura, T., & Sanefuji, W. (2020). The influence of parental rearing styles on university students’ critical thinking dispositions: The mediating role of self-esteem. *Thinking Skills and Creativity*, 37, 100679. <https://doi.org/10.1016/j.tsc.2020.100679>
- Wolfradt, U., Hempel, S., & Miles, J. N. V. (2003). Perceived parenting styles, depersonalisation, anxiety and coping behaviour in adolescents. *Personality and Individual Differences*, 34(3), 521–532. [https://doi.org/10.1016/S0191-8869\(02\)00092-2](https://doi.org/10.1016/S0191-8869(02)00092-2)
- Xu, J., Tong, M., Gao, Q., Xu, Y., Chen, J., & Long, T. (2026). The impact of project-based learning (PjBL) on creative thinking in fourth graders: A quasi-experimental study in blended science synchronous classroom. *Thinking Skills and Creativity*, 61, Article 102165. <https://doi.org/10.1016/j.tsc.2026.102165>
- Yalçın, H., & Özdemir, E. (2021). Anne-baba tutumlarının ergenlerdeki depresyon, anksiyete ve stres düzeylerine etkisi [The effect of parental attitudes on levels of depression, anxiety, and stress among adolescents]. *Karatay Sosyal Araştırmalar Dergisi*, 7, 229–248. <https://izlik.org/JA82GK38RJ>
- Yan, Z., Yu, S., & Lin, W. (2024). Parents’ perceived social support and children’s mental health: The chain mediating role of parental marital quality and parent-child relationships. *Current Psychology*, 43(5), 4198–4210. <https://doi.org/10.1007/s12144-023-04625-x>
- Yetgin, A. & Katrancı, M. (2020). İlkokul öğrencilerinin eleştirel düşünme becerilerinin çeşitli değişkenler açısından değerlendirilmesi [Evaluation of primary school students’ critical thinking skills in terms of various variables]. *International Journal of Languages’ Education and Teaching*, 8(4), 81–91. <https://doi.org/10.29228/ijlet.46576>
- Yıldırım Döner, S., & Demir, S. (2022). Developing the critical thinking disposition scale for secondary school students: A validity and reliability study. *Pamukkale University Journal of Education*, (54), 99–129. <https://doi.org/10.9779.pauefd.823427>
- Yılmaz, A. (2025). Tasarım tabanlı araştırma yöntemi ile hazırlanan fen etkinliklerinin ortaokul öğrencilerinin yaratıcı ve eleştirel düşünme becerilerine etkisi: Sınıf içi deneysel uygulama çalışması [The effect of science activities designed through the design-based research method on middle school students’ creative and critical thinking skills: a classroom-based experimental study]. *Kastamonu Education Journal*, 33(3), 572–584. <https://doi.org/10.24106/kefdergi.1748663>
- Yurt, E., & Kara, M. (2025). The influence of parental attitude, school climate, and classroom management on critical thinking: The significance of inclusive classrooms in higher education. *SAGE Open*, 15(3). <https://doi.org/10.1177/21582440251382061>
- Zeevi-Cousin, A., & Lavenda, O. (2023). The mediating role of parenting style in the relationship between parents’ openness to different ways of thinking and child anxiety. *Children*, 10(9), 1564. <https://doi.org/10.3390/children10091564>
- Zhao, X., & Yang, J. (2021). Fostering creative thinking in the family: The importance of parenting styles. *Thinking Skills and Creativity*, 41, 100920. <https://doi.org/10.1016/j.tsc.2021.100920>



This page is intentionally left blank.
www.iejee.com

Integrating Water Sustainability into Primary Science Education: Teachers' Pedagogical Rationales, Opportunities, and Barriers

Puji Winarti^a, Sri Haryani^{b*}, Sri Wardani^c, Langlang Handayani^d

Received : 3 March 2026
 Revised : 25 June 2026
 Accepted : 30 September 2026
 DOI : 10.26822/iejee.2026.462

^a Puji Winarti, Universitas Negeri Semarang, Semarang, Indonesia, Darul Ulum Islamic Centre Sudirman University, Semarang, Indonesia.
 E-mail: pujiwinartirulian@gmail.com
 ORCID: <https://orcid.org/0000-0002-3292-7164>

^{b*} **Corresponding Author:** Sri Haryani, Universitas Negeri Semarang, Semarang, Indonesia.
 E-mail: haryanikimia83@mail.unnes.ac.id
 ORCID: <https://orcid.org/0000-0002-7699-326X>

^c Sri Wardani, Universitas Negeri Semarang, Semarang, Indonesia.
 E-mail: menuksriwardani@mail.unnes.ac.id
 ORCID: <https://orcid.org/0000-0002-5935-0449>

^d Langlang Handayani, Universitas Negeri Semarang, Semarang, Indonesia.
 E-mail: langlanghandayani@mail.unnes.ac.id
 ORCID: <https://orcid.org/0000-0003-2131-5370>

Abstract

Water sustainability is a key global concern highlighted in SDG 6 (Clean Water and Sanitation). Integrating this topic into primary science education is strategic for fostering scientific literacy, environmental awareness, and responsible values from an early age. This study explores primary school teachers' pedagogical perspectives on the importance of water sustainability, opportunities for integration, and perceived barriers to classroom implementation. A qualitative descriptive approach was employed using an open-ended online questionnaire completed by 141 primary school teachers across Indonesia. The data were analyzed thematically using NVivo. The findings reveal five rationales for integration: environmental awareness, science literacy, character and values education, alignment with the SDGs, and relevance to students' daily lives. Teachers identified integration opportunities through science topics, environmentally themed learning, and project-based activities. However, teachers anticipated barriers including limited resources, low confidence, dense curriculum demands, time constraints, local environmental conditions, and challenges in selecting appropriate instructional methods. The novelty of this study lies in examining primary school teachers' pedagogical reasoning regarding water sustainability and identifying opportunities and barriers to its integration into science learning. The findings suggest that the integration of water sustainability is influenced by curriculum demands, teacher preparedness, available resources, and local environmental conditions. This study contributes to sustainability education research by providing insights into how sustainability goals are interpreted and implemented in primary school contexts.

Keywords:

Primary science education; Water sustainability; Teachers' pedagogical reasoning; Sustainability education; Science education.



www.iejee.com
 ISSN: 1307-9298

2026 Published by KURA Education & Publishing.
 This is an open access article under the CC BY-NC-ND license. (<https://creativecommons.org/licenses/by/4.0/>)

Introduction

Water and its sustainability have become crucial issues in global development. Sustainable Development Goal (SDG) 6, clean water and sanitation, explicitly addresses water management and access. Water is also intrinsically linked to several other SDGs, such as SDG 3 (Good Health

and Well-being), SDG 13 (Climate Action), SDG 14–15 (Life Below Water and Life on Land), as well as socio-economic issues, including poverty and food security. This highlights that water is a cross-cutting theme within sustainability development frameworks. Water sustainability refers to the conservation and equitable accessibility of clean water to support human consumption, agricultural practices, and biodiversity for future generations (Haque et al., 2025; Ramírez-Segado et al., 2025).

In the context of primary education, integrating water sustainability topics is a strategic step to foster sustainability awareness from an early age and to develop environmentally responsible citizens. Various studies have shown that the primary school age is a crucial period for shaping students' values, attitudes, and environmental awareness (Irimia et al., 2025; Praveena & Themudu, 2022; Ruiz-Garzón et al., 2021). Recent studies have demonstrated that teaching these topics in primary school can help students comprehend the intricate relationships between human needs, the environment, and public policy (Pozo-Muñoz et al., 2023; Tsiaras & Chavenetidou, 2024).

The urgency of teaching water sustainability is particularly relevant in Indonesia, as water remains a pressing environmental and social issue. National environmental reports indicate that it is facing increasing water-related challenges. The Minister of Environment and Forestry, Republic of Indonesia (2023), reports that river water quality in many regions does not meet established standards, with biological oxygen demand (BOD), chemical oxygen demand (COD), fecal coliform, and total coliform levels exceeding permissible thresholds in most provinces due to domestic pollution. BPS-Statistics Indonesia (2025) also reported disparities in household access to safe drinking water between regions, with provinces such as Papua and Bengkulu having lower levels of access compared to the national average. These conditions highlight the importance of integrating water-related sustainability issues into primary science education to foster environmental awareness and responsible citizenship among young learners.

In this context, water sustainability is more than an environmental or moral concern (Blom & Karrow, 2024). It provides a meaningful context for developing water awareness. Awareness of water sustainability cannot be developed solely through cognitive learning, but needs to be embedded in learning experiences that encourage critical reflection, empathy, and active engagement with real-world problems (Mostacedo-Marasovic et al., 2022). Teachers are central to this process because they guide students in interpreting water-related sustainability issues through scientific reasoning. Consequently, teachers' beliefs,

perspectives, and preparedness strongly influence how these topics are addressed in classroom practice (Boon, 2024).

However, integrating water sustainability into science learning is not a simple task. Teachers' capacity to integrate these topics is influenced by their pedagogical content knowledge, confidence, and access to contextual teaching resources (Forsler et al., 2024). Many teachers still struggle to translate broad sustainability goals into classroom practices that are appropriate to students' age and characteristics, particularly when curricular guidance and professional preparation remain limited (Forsler et al., 2025; Uygun & Bayrakci, 2025). For this reason, it is important to examine how teachers perceive the urgency of water sustainability, as well as the opportunities and challenges they encounter when integrating it into science learning.

Theoretical framework

This research is based on the perspectives of ecopedagogy and Education for Sustainable Development (ESD). Ecopedagogy views education as a space for building ecological awareness, ethical responsibility, and a critical understanding of the relationship between humans and the natural environment (Kahn, 2010; Misiaszek, 2015). From this perspective, water sustainability is viewed as both a scientific topic on natural resources and an ecological, social, and pedagogical issue. Ecopedagogy emphasizes the importance of helping students critically understand how everyday human activities are connected to water sustainability. This perspective aligns with the goals of ESD, which emphasize that education should be oriented toward mastering knowledge, developing values, attitudes, and actions that support sustainable living (UNESCO, 2020).

In the context of primary science education, the issue of water provides a learning context closely related to students' everyday experiences. Topics such as water use, environmental cleanliness, the water cycle, natural resources, pollution, and conservation help students connect scientific concepts to the realities of their surroundings. Therefore, science learning oriented toward water sustainability can foster awareness, responsibility, and critical reflection on environmental issues that impact human life and ecosystems (Cabarcas-Fuentes & Rodríguez-Chueca, 2026; McCarroll & Hamann, 2020).

This study uses the concept of teachers' pedagogical reasoning as the basis for analysis. Pedagogical reasoning refers to the professional process of teachers in understanding material, transforming it into a teachable form, selecting appropriate learning strategies, and considering student characteristics and the learning context (Elliott, 1996; Shulman,

1987). Within this framework, teachers' pedagogical reasoning is used to analyze how primary school teachers interpret water sustainability as a learning issue, provide pedagogical reasons for integrating it into science learning, and identify opportunities and barriers to classroom implementation. By integrating perspectives from ecopedagogy, ESD, and teachers' pedagogical reasoning, this study views water sustainability education as both a scientific and a pedagogical issue.

Although ESD and SDG 6 have been widely emphasized in global education research, empirical studies focusing on primary school teachers' perceptions of water sustainability remain limited. In Indonesia, although there is no national education policy or school curriculum that explicitly addresses water sustainability at the primary level, the broader direction of education policy provides space for integrating environmental and sustainability issues. This is reflected in the emphasis on scientific literacy, character education, and contextual learning using real-world issues. These policy orientations align with the principles of ESD and support the integration of global sustainability goals into national educational practice. Within this framework, water sustainability can be positioned as a relevant topic in primary education, with teachers playing a key role in translating global goals into classroom practice.

Most existing studies on water sustainability education have focused on students' understanding of water-related issues (Pozo-Muñoz et al., 2023; Ramírez-Segado et al., 2021; Zhan et al., 2019), the development of sustainability-related competencies and skills (Garcia et al., 2025; Pozo-Muñoz et al., 2023; Rey et al., 2022; Tsiaras & Chavenetidou, 2024) and instructional strategies (Imaduddin & Eilks, 2024; Scarborough et al., 2022; Stephenson Reaves, 2023). In contrast, teachers' pedagogical reasoning, contextual interpretations, and challenges in implementing water sustainability have received far less attention. This gap is important because teachers play a central role in translating sustainability goals into meaningful learning experiences through their selection of teaching materials, instructional approaches, and the integration of contextual environmental issues into primary science classrooms (Green & Somerville, 2015; Inoue et al., 2016; Kwee & Dos Santos, 2023). Without understanding teachers' pedagogical reasoning, it remains difficult to explain how water sustainability can be meaningfully and realistically integrated into primary science classrooms.

Furthermore, most prior research was conducted in Western or secondary education contexts. This leaves a knowledge gap regarding how teachers in developing countries, particularly in Indonesia, conceptualize and integrate water sustainability in classroom teaching.

Unlike previous studies that primarily emphasized student outcomes and instructional approaches, this study examines primary school teachers' pedagogical reasoning, perceived opportunities, and obstacles in teaching water sustainability. These findings provide insights to inform sustainability-oriented pedagogy and teacher education programs at the primary level.

This study aims to explore primary school teachers' pedagogical perspectives on integrating water sustainability into primary science education, including their perceptions of its importance, the opportunities they identify for integration, and the barriers they anticipate in classroom implementation. Therefore, this study was guided by the following research questions:

1. How do primary school teachers perceive the importance of integrating water sustainability into primary science education?
2. What opportunities do primary school teachers identify in integrating water sustainability into primary science learning?
3. What barriers do primary school teachers anticipate in implementing water sustainability education in primary classroom practice?

Method

Research Design

This study employs a qualitative descriptive approach to explore primary school teachers' perspectives on water sustainability in primary science (Mackiewicz, 2018). This design was chosen because the study aimed to obtain a broad and practice-oriented understanding of teachers' rationales, perceived opportunities, and anticipated barriers, rather than to examine lived experiences in depth or analyze a bounded instructional setting.

Unlike phenomenological studies, which focus on the essence of individual experiences or case studies examining specific contexts or cases, a qualitative descriptive approach was considered more appropriate for capturing diverse perspectives across teachers from different regions. The use of an open-ended online questionnaire further supported this aim by enabling the collection of varied reflections from a wider participant pool, aligned with the exploratory nature of the research questions. The use of written open-ended responses allowed participants to explain their experiences and perspectives in their own words and at their own pace. The online format also enabled teachers from different geographical regions to participate, supporting the exploratory aim of capturing varied pedagogical perspectives on water sustainability.

Participants

Participants included 141 primary school teachers (33 male and 108 female) from diverse regions across Indonesia. Teachers were recruited using voluntary response sampling through an open online invitation distributed via accessible professional teacher networks. This sampling approach was convenience-based rather than purposive, as participation depended on teachers' self-selection. While this strategy enabled broad geographical coverage, it may have favored teachers with a greater interest in sustainability-related topics. To broaden participation, the invitation was disseminated across multiple teacher communities representing different regions and socio-economic contexts.

Table 1.
Demographic characteristics of participants

Characteristic	Category	n
Gender	Male	33
	Female	108
Teaching experience	1–5 years	60
	6–10 years	42
	years	22
	>15 years	17
Region of employment	Java	42
	Sumatra	30
	Kalimantan	31
	Eastern Indonesia	38
School type	Public primary school	98
	Private primary school	43

Table 1 presents the demographic characteristics of the participants. In this study, Eastern Indonesia refers to provinces in Sulawesi, Nusa Tenggara, Maluku, and Papua. Given the sampling strategy and study aims, the findings are intended to provide exploratory insights into teachers' rationales, perceived opportunities, and barriers rather than to represent the broader teacher population. Although the study relied on open-ended online responses, many participants provided relatively detailed explanations and reflections regarding water sustainability in primary science education. Responses from teachers across different regions showed recurring patterns and variations in perspectives, which contributed to the development of themes during the analysis. This diversity of responses supported the study's exploratory aim of capturing a broad range of teachers' pedagogical perspectives.

Data Collection

The open-ended questionnaire was developed based on the study's research focus and the concept of teachers' pedagogical reasoning, particularly teachers' considerations regarding instructional goals, classroom integration and contextual challenges in

teaching water sustainability. The questionnaire was then organized into three main sections: (1) teachers' views on the importance and rationale for teaching water sustainability in primary schools; (2) perceived opportunities for integrating the topic into science learning; and (3) anticipated obstacles or challenges in implementation.

To ensure the instrument's content validity and clarity, the initial draft was reviewed by two experts in science education and sustainability. Their feedback led to minor revisions in wording and sequencing to improve contextual relevance. The revised questionnaire was piloted with six teachers from a primary school, and further refinements were made based on their feedback to improve clarity and comprehensibility.

The open-ended online questionnaire was distributed electronically and remained accessible for four weeks (February–March 2025). Before completing the questionnaire, participants were informed of the study's purpose, the voluntary nature of participation, and their right to withdraw at any time. Informed consent was obtained electronically prior to participation. No personally identifiable information was collected, and participants' responses were anonymized during analysis. All data were stored securely and were accessible only to the research team. Because the study involved voluntary participation, anonymous responses, and no sensitive personal data, formal ethical approval was not required under institutional research guidelines.

The questionnaire consisted of open-ended questions designed to elicit teachers' perspectives on water sustainability. Example items included: (1) Why do you think water sustainability is important to be taught in primary school? (2) In what ways do you see opportunities to integrate water sustainability into primary school learning? (3) What challenges do you anticipate when implementing water sustainability topics in your classroom?

Data Analysis

Qualitative responses were analyzed using thematic analysis following Braun and Clarke's six-phase framework (Braun & Clarke, 2006). The analysis also drew on later developments in reflexive thematic analysis that emphasize the interpretive role of researchers in theme construction (Braun et al., 2022). Coding was conducted primarily inductively, meaning that themes were developed from participants' responses while still being guided by the study's focus on teachers' pedagogical reasoning. The analysis focused mainly on semantic themes, emphasizing participants' explicit perspectives and explanations regarding water sustainability in primary science education. All responses were read several times to develop familiarity with the data. Initial open coding

was used to identify recurring ideas, rationales, and perceptions. Codes were then reviewed, refined, and grouped into broader themes through an iterative process of comparison and interpretation. Theme development continued through iterative comparison and refinement until the themes adequately represented recurring patterns in the dataset. A single response could be assigned multiple codes when it reflected more than one idea or perspective. NVivo software was used to organize responses, manage coding categories, retrieve coded excerpts, and facilitate comparisons during theme development and refinement. The findings section presents representative excerpts and coding categories to illustrate how themes were developed from participants' responses and to enhance analytical transparency. Additional examples of the coding and theme development process are also provided in Appendix A.

To enhance credibility and trustworthiness, the coding process involved the first researcher and three additional researchers with experience in science education and qualitative research. After the first researcher completed initial open coding, the coding results and emerging themes were collaboratively reviewed through a series of peer debriefing discussions. During these discussions, differences in interpretation were examined by revisiting participants' responses, comparing coding rationales, and refining theme boundaries until agreement was reached. An audit trail was maintained throughout the analytical process by documenting coding revisions, theme development procedures, analytical notes, and interpretive decisions. In addition, member checking was conducted with six participants, who were asked to review summaries of the preliminary themes to verify whether the interpretations accurately reflected their perspectives. Reflexive memos were also written throughout the analysis to monitor researchers' assumptions, interpretive decisions, and potential biases. The research team's background in science education and sustainability-related research informed the study's interest in exploring teachers' perspectives on water sustainability. Throughout the analysis, the research team remained aware of the potential influence of these perspectives on data interpretation and addressed potential biases through reflexive discussions and memo writing. These procedures helped strengthen the credibility, dependability, and confirmability of the findings.

Findings

Primary Teachers' Pedagogical Rationales for Integrating Water Sustainability into Primary Science Education

This study reveals the various reasons primary school teachers cite for integrating water sustainability issues

into primary school science learning. Based on the results of the thematic analysis, five main themes were obtained that reflect the participants' thinking frameworks. Each theme is supported by narrative quotes and frequency data that demonstrate the strength of representation for each reason, as provided in Table 2 below.

Table 2 presents the themes and codes related to teachers' rationales for integrating water sustainability in primary science. In this study, frequency counts (f) indicate the recurrence of specific codes within each theme and are used to illustrate the relative prominence of themes in the dataset, rather than to quantify responses at the theme level.

The first theme emerging from teachers' responses was raising environmental awareness. Teachers viewed the introduction of water sustainability issues from an early age as important for helping students develop awareness of the need to protect water resources. Their responses also linked water sustainability learning with everyday habits, such as using water wisely and caring for the environment from an early age. Many teachers emphasized the importance of nurturing a generation that is more environmentally aware through learning activities connected to students' daily lives.

Another theme concerned the relationship between water sustainability and the SDGs, particularly issues related to clean water access and sustainable water resources. Teachers recognized that water issues extend beyond the science content taught in the classroom. Many teachers also viewed water-related learning as a way to introduce students to broader environmental issues. This suggests that many teachers were already concerned about future environmental sustainability. Teachers also viewed water sustainability as part of character and values education. In their responses, water-related issues were often associated with developing habits of wise water use, gratitude for natural resources, and responsibility toward the environment. Teachers also connected water sustainability learning with the development of students' everyday attitudes and habits.

In addition, relevance to daily life emerged as one of the reasons why this topic was considered suitable for primary science learning. Teachers referred to students' routine activities, such as bathing, washing, drinking, and watering plants, as contexts closely related to their everyday experiences. This connection to daily activities made water sustainability easier for students to understand because it was closely linked to situations they encounter in everyday life. The final theme related to strengthening scientific literacy. Teachers viewed water as an important concept for understanding its relationship with living things and

the environment. Teachers also considered water sustainability helpful for introducing science concepts related to the environment and daily life.

Primary Teachers' Perceived Opportunities for Integrating Water Sustainability into Primary Science Education

Teachers described various opportunities for integrating water sustainability into primary school learning. Their responses showed that water sustainability could be connected to science topics, environmental themes across subjects, and contextual learning activities. An overview of the main themes identified from participants' responses is presented in Table 3.

The most frequently mentioned opportunity concerned integrating water sustainability into existing primary science topics across grade levels. Teachers commonly associate water sustainability with topics such as ecosystems, natural resources, environmental change, and daily life contexts. Several participants also described more specific examples of how water sustainability could be integrated into science learning from Grades 1 to 6. Representative examples provided by teachers are presented in Table 4.

The examples presented in Table 4 illustrate how teachers connected water sustainability with science

topics across different grade levels. These responses suggest that teachers viewed water sustainability as relevant to a wide range of primary science contexts and everyday environmental issues.

The second opportunity identified from teachers' responses was the integration of water sustainability through environmental themes across subjects. Teachers associated water sustainability with environmental topics discussed in several school subjects, such as citizenship education, language learning, mathematics, arts, and social science. Examples provided by teachers are presented in Table 5.

The examples in Table 5 illustrate that teachers associated water sustainability with environmental themes across different subjects. Water-related issues were linked with topics such as environmental responsibility, natural resources, and environmental change in various learning areas.

The third opportunity identified from teachers' responses was the integration of water sustainability through thematic or end-of-semester projects. Teachers described project activities as opportunities for students to explore water-related issues in more contextual ways. Representative examples provided by teachers are presented in Table 6.

Table 2.

Rationale for water sustainability in primary science

Theme	Code	Example statement	f
Raising environmental awareness	Early water awareness	It is very important because the earlier it is taught to children, the more they will understand the importance of water for the sustainability of life (T-5)	35
	Eco-conscious generation	Knowledge about the importance of water and how to save it from an early age will form good habits that will last into adulthood (T-25)	18
Alignment with SDGs	Availability of clean water	It is very important to teach this because children need to understand the importance of sustainable clean water (T-4)	6
	Water sustainability	Water sustainability is very important for primary school students because it is for life now and in the future (T-6)	19
Character and values education	Good and being grateful	So that students really know the benefits of using water and continue to be grateful (T-1)	1
	Use water wisely	It is very important that children understand how to use water wisely (T-14)	23
	Build positive character	Develop good habits in using water (T-8)	9
	Environmental stewardship	Students can become more environmentally conscious and can help create a more sustainable future (T-13)	12
Relevance to daily life	Bathing	This topic is very relevant to their daily lives, like bathing.....(T-7)	8
	Drinking	...water is a basic necessity in daily life, such as for drinking, cooking, etc (T-3)	4
	Washing	Students must know how to maintain clean water that can be used daily, such as, washing clothes.....(T-15)	5
	Watering plants	Water is very necessary for humans in daily life, also for animals, and for watering plants (T-16)	2
Supporting science literacy	Water is a source of life	Because water is a source of life for living creatures, from the beginning of life until death (T-31)	43
	Importance of the ecosystem balance	Important to maintain the balance of the ecosystem (T-1)	9

Table 3.*Overview of teachers' perceived opportunities for integrating water sustainability into primary school learning*

Theme	Description	f
Integration into primary science topics	Water sustainability can be integrated into science topics from Grades 1–6	111
Cross-subject integration through environmental themes	Water sustainability can be connected with environmental themes across subjects	25
Integration through thematic projects	Water sustainability can be integrated through thematic or end-of-semester projects	10

Table 4.*Representative examples of water sustainability integration across primary science topics*

Grade	Primary Science Topic	Example Statement
I	Natural events	"...in the theme of introducing natural disasters such as floods and droughts..." (T-99)
II	Clean and healthy environment	"The topic of water sustainability is very flexible and can be integrated into various science lesson themes, including the theme Clean and Healthy Environment (Grade 2 Primary School)" (T-39)
III	Ecosystem	"This topic is integrated with the material on the water cycle, natural resources, and ecosystems" (T-42)
IV	Natural resources and environment	"Science: Natural Resources. Because it also includes how to utilize natural resources, including the water cycle" (T-25)
V	Harmony in the ecosystem	"The role of water in maintaining ecosystem balance and the impact of water pollution (grade 5)" (T-67)
VI	Our Earth is in danger	"For grade 6, the topic can be connected with the chapter 'Our Earth is Threatened with Danger'" (T-26)

Table 5.*Examples of water sustainability integration through environmental themes across subjects*

Subjects	Environmental theme	Example statement
Pancasila and civic education	Citizens' responsibility in protecting natural resources	Activity: Discussion of rights and obligations in water use (T-67)
Indonesian language	Write an appeal text, create a poem, or a story (about the importance of water)	Making water-saving campaign posters, composing speeches about water (T-81)
Mathematics	Calculating volume (of liquids)	The project records daily water usage at home (T-92)
Arts, Culture and Crafts	Creating artwork illustrating water conservation messages (comparisons of clean and polluted rivers, or campaigns encouraging responsible water use).	Painting about clean and polluted rivers (T-94)
Social science	Sustainable development, environmental change	Discuss the social impacts of water shortages or water management (T-49)

Table 6.*Examples of water sustainability integration through thematic projects*

Project Activity	Example Statement
Water-saving campaign project	"Students are asked to make posters about keeping water, rivers, and the sea clean" (T-81)
Observation of water use in daily life	"Students keep a simple weekly record of how they use water wisely at home and at school during the semester" (T-55)
Environmental cleanliness project	"Students can join activities such as cleaning rivers, water springs, or nearby wetlands" (T-67)

The examples in Table 6 suggest that teachers preferred simple, contextually relevant project activities closely related to students' daily lives and local environments.

Teachers also identified opportunities to integrate water sustainability into learning methods and approaches suitable for primary science learning. Many teachers emphasized hands-on activities, such as observing water use in daily life, visiting nearby water sources, and creating simple tools related to water recycling. Other responses highlighted the use of simple experiments, including water filtration and evaporation activities. Teachers viewed these activities as helpful for understanding science concepts through direct experience. They also mentioned using videos and picture stories, particularly for lower-grade students. These media were considered helpful for explaining environmental issues in more engaging ways. In addition, classroom discussions were viewed as opportunities for students to share experiences and express their opinions about water-related issues. Several teachers also referred to project-based learning, problem-based learning, and contextual learning approaches. These responses suggest that teachers preferred learning activities closely connected to students' everyday experiences and surroundings.

Primary Teachers' Perceived Barriers to Integrating Water Sustainability into Primary Science Education

Based on primary school teachers' perceptions of the obstacles they might encounter when integrating water sustainability into science learning, six relevant themes were identified, as presented in Table 7.

In this table, the frequency (f) represents the number of teachers who mentioned each obstacle theme in their responses, indicating the relative prominence of obstacles across participants. The obstacles described above arise from various internal and external factors that can influence the effectiveness of integrating water sustainability into science learning in primary schools.

The first theme that emerged from the teachers' responses related to limited school facilities and resources. Teachers cited a lack of learning media, limited laboratories, and limited facilities for environmental observation as obstacles to teaching water sustainability. Some teachers felt that learning about water would be easier for students to understand if supported by practical activities or direct observation. However, school conditions did not always allow for this. One teacher explained that their school "did not yet have sufficient facilities to support direct environmental observation activities" (T-9). This response indicates that some teachers still view water sustainability as material that requires special learning facilities, adequate media or specific practical activities to facilitate student understanding.

Teachers also mentioned limited understanding and confidence in integrating water sustainability issues into science lessons. Several responses indicated that teachers still struggle to grasp the concept of water sustainability in depth and connect it to primary school learning materials. One teacher stated that "the concept of water sustainability is still quite difficult to explain to primary school students" (T-24). Furthermore, some teachers also found it difficult to explain the topic in a way that was simple, engaging, and appropriate to the characteristics of primary school students. Teacher responses indicated that they were still confused about how to explain water sustainability issues in a way that aligns with primary school students' understanding.

Another frequently encountered obstacle relates to curriculum limitations and learning time. Teachers described how the dense learning materials and demands to meet curriculum targets limited their time for in-depth discussion of additional topics, such as water sustainability. Some teachers also expressed hesitation in integrating the topic due to a lack of clear emphasis in the curriculum documents. One teacher stated that "there is already quite a lot of material to cover in school, making it difficult to add in-depth discussion" (T-11). This response indicates that some teachers still view water sustainability as an additional topic outside of the core material to be covered in class.

Table 7.

Obstacles to integrating water sustainability into science learning in primary schools

Theme	f	Statement example
Limited resources and facilities	66	Lack of supporting facilities at school (T-7)
Limited teacher knowledge and confidence	26	Lack of in-depth understanding of the concept of water sustainability among teachers (T-5)
Curriculum constraints	25	Dense Curriculum Focus (T-4)
Environmental context	25	Students tend to underestimate the benefits of water because there is an abundance of water at home (T-30)
Time limitation	20	limited learning time at school (T-16)
Method uses	18	The teacher has not been able to package the explanation regarding this topic to students in a good and interesting way, so that students can understand and accept the explanation from the teacher (T-70)

Teacher responses also demonstrated that environmental conditions influence how students perceive water sustainability issues. In some areas where students live, water is abundant, leading them to believe it will always be available. Teachers believe this makes students less aware of the importance of water conservation. One teacher explained that “students are accustomed to seeing water available in abundance, so they feel it will always be available” (T-81). In contrast, teachers in areas with sanitation issues or limited water resources described water issues as more closely related to students’ everyday experiences. This difference in context suggests that students’ environmental experiences shape how they understand water sustainability issues in their learning.

Furthermore, limited learning time presents a challenge in implementing more contextual learning. Teachers feel that activities such as projects, field observations, or discussions require more time than regular learning. Class hours in primary schools are relatively limited. Several teachers reported that school time is often insufficient to implement experiential learning activities. Teachers also reported difficulty selecting learning methods appropriate to students’ characteristics, classroom conditions, and school facilities. One teacher explained that they “seem to have difficulty explaining the topic of water sustainability in a way that is interesting and easy for students to understand” (T-78). Teacher responses indicate that they still struggle to adapt water sustainability learning to classroom conditions and the characteristics of primary school students.

Discussion

The findings of this study indicate that primary school teachers have diverse reasons for emphasizing the importance of integrating water sustainability into science learning. Teachers view water sustainability as a topic close to students’ daily lives, making it relevant for classroom learning. The emphasis on awareness and character building indicates that water sustainability is understood more as part of the development of students’ behavior and habits than as mastery of scientific concepts. Teachers also frequently link learning about water with wise water use, gratitude, and environmental responsibility. This emphasis on the formation of sustainable values, awareness, and behavior at the primary school level is also widely found in previous research (Echegoyen & Martín, 2021; Siraj & Pramling, 2016; Višnjić-Jevtić et al., 2022; Wang et al., 2024). This condition shows that sustainability learning at the primary school level is still more often positioned as a means of forming awareness and behavior rather than as a space to understand the complexity of sustainability issues in more depth (Calikusu & Sabanci, 2025)

Although teachers have linked water sustainability to the SDGs, this understanding tends to focus on the importance of clean water and the responsibility to protect the environment. Teachers have not yet linked water issues to other sustainability issues, such as health, climate change, or ecosystem sustainability. Yet, water issues are closely linked to various global sustainability issues (UN-Water, 2016). The SDGs are not yet fully understood as a learning framework that interconnects sustainability issues. Teachers still tend to understand the SDGs separately and unrelatedly. This finding supports previous research showing that teachers often struggle to understand and integrate the SDGs more holistically into their learning (Ampartzaki et al., 2021; Hashim, 2023; Imaduddin & Eilks, 2024). However, in the ESD concept, the SDGs are expected to serve as a basis for developing integrated learning strategies.

Teachers also identified concrete entry points for learning. They cited a variety of science topics from grades 1 to 6 as relevant contexts. This was possible because they believed that integrating water sustainability into core science themes was a meaningful form of learning (Tsiaras & Chavenetidou, 2024). Teachers view cross-subject integration and end-of-semester thematic projects as learning opportunities, for example through observing water use in students’ environments (T-39; T-67). This kind of cross-disciplinary approach has proven effective in preparing students to face complex and interconnected sustainability challenges, including those related to water (Romulo et al., 2024). In addition, integration into an end-of-semester thematic project with adequate time can increase student engagement, develop critical skills, and strengthen environmental awareness (Valenzuela-Morales et al., 2022; Yang, 2021). This project-based orientation is also supported by empirical evidence showing that sustainability-based water projects can be implemented through an inquiry-based approach. Simple, low-cost water purification activities are one example of a project context with limited resources (Fikri et al., 2026). The identified opportunities indicate that teachers are beginning to consider pedagogical factors in integrating water sustainability into learning.

Regarding implementation barriers, this study revealed a clear contradiction between awareness and implementation. On the one hand, teachers demonstrated strong awareness and clear pedagogical reasoning in integrating water sustainability into classroom learning. However, teachers also predicted classroom implementation barriers. Teachers’ awareness of the importance of water sustainability was not always accompanied by pedagogical readiness and structural support in schools. Teachers recognized that water sustainability issues were important to teach, but they still faced

limitations in translating these issues into practical learning practices. Various barriers stemmed from limited facilities and resources, including technology and costs (Espinoza-Ramos & Capucci-Polzin, 2025; Rosmiati et al., 2025), low teacher knowledge and confidence (Alsina & Vásquez, 2025), pressure from a dense curriculum and limited learning time (Do & Thinh, 2025), and the influence of environmental context on the perception of the urgency of water issues (Hamwy et al., 2023), also experienced by other teachers in implementing continuous learning. Difficulties in selecting appropriate learning methods also indicate the need for flexible time planning and adequate pedagogical support (Hamwy et al., 2023). This contradiction between intention and practice reflects the gap between teachers' normative commitments and their pedagogical capacities, as also reported in various international studies on sustainability education (Alsina & Vásquez, 2025).

In the context of Indonesian education, the contradiction between awareness and implementation needs to be understood as a structural issue. Thematic learning in character education offers opportunities to integrate water sustainability across subjects. However, this differs from several countries that have explicitly incorporated sustainability and the SDGs into science curriculum standards (Annamalai et al., 2025; Ferrer-Est & Chalmers, 2021), integration in Indonesia still relies heavily on teacher interpretation and initiative. Furthermore, local environmental conditions also influence perceptions of the urgency of water issues. In areas with abundant water, water sustainability is often seen as less pressing. In areas with limited water, however, the issue is more easily contextualized. This reliance on these factors can potentially lead to uneven integration practices across schools and contexts. As a result, water sustainability learning develops more based on local needs and experiences than through systematic planning. The interplay between structural constraints and social contexts means that water sustainability has not yet become a consistent framework in primary school learning. Without explicit curricular and pedagogical guidance, sustainability learning often relies on individual teachers and varies across schools (Vesterinen, 2024).

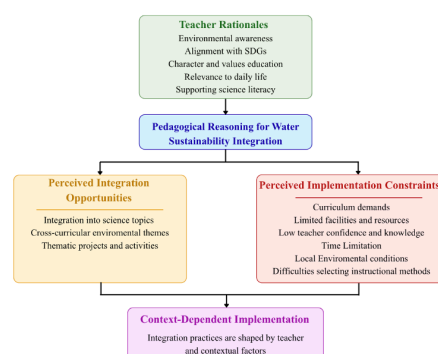
Water sustainability learning in Indonesia is also shaped by everyday practices and community values. Teachers often use domestic activities such as bathing, washing, and household water use as entry points for learning because they are close to students' lives. Water sustainability is not positioned as a scientific topic, but as a form of social and moral responsibility. This sociocultural context explains why teachers often link water sustainability learning with character education and attitude development rather than focusing solely on scientific concepts.

Overall, the research findings indicate that teachers view water sustainability as an issue close to students' daily lives and relevant to primary school learning. They emphasize the importance of environmental awareness and consider how water issues can be connected to science topics, learning projects, and classroom activities. However, these integration efforts are still constrained by various limitations, including curriculum, learning time, facilities, and teacher preparedness. As a result, the implementation of water sustainability learning varies among teachers and across school contexts. This research suggests that integrating water sustainability in primary schools depends on teacher awareness, pedagogical support, and the learning conditions within the school.

The findings also suggest that teachers' perspectives on water sustainability are shaped by environmental concerns and practical classroom considerations. Although teachers recognized the importance of water sustainability and identified various opportunities for integration, their implementation decisions were influenced by curriculum demands, available resources, student characteristics, and local environmental conditions. These findings indicate that integrating water sustainability into primary science learning involves balancing sustainability goals with the realities of classroom practice. This research also contributes to ESD studies. The findings suggest that teachers interpret water sustainability as an environmental topic and issue closely related to students' daily experiences, character development, and classroom learning practices. This indicates that sustainability integration in primary education is influenced by sustainability goals, teachers' pedagogical considerations and the realities of classroom contexts.

Based on the findings and discussion, Figure 1 presents a conceptual model of the relationships among teachers' rationales, integration opportunities, implementation constraints, and variations in water sustainability learning practices in primary schools.

Figure 1.
Conceptual model of teachers' pedagogical reasoning for integrating water sustainability into primary science education.



Conclusion

This study confirms that primary school teachers have a strong awareness of the importance of water sustainability and can identify various entry points for integrating it into science learning. However, this integration remains constrained by curriculum pressures, resource constraints, and teachers' pedagogical preparedness, resulting in often situational and individual-driven learning practices. This research contributes to the study of ESD by showing that teachers' reasoning about water sustainability reflects both sustainability-related values and practical considerations related to curriculum and classroom implementation. These findings underscore the need for more explicit curricular-pedagogical frameworks, contextualized teacher professional development, and realistic and adaptive learning models to support more systematic integration of water sustainability in primary schools.

Practical Implications

Based on the findings, several practical implications can be identified. First, primary school teachers can use existing science topics as entry points for integrating water sustainability without adding new curriculum content. Second, schools and curriculum developers should provide simple and flexible learning tools, such as concise lesson plans, observation activities, and locally based thematic projects that can be implemented with limited facilities. Third, teacher professional development should emphasize contextual pedagogical strategies, including local case-based training, microteaching, and the sharing of best practices. Fourth, support from schools and local governments is needed to enable pedagogical flexibility through cross-subject projects and collaboration with local communities, ensuring that water sustainability integration does not rely solely on individual teacher initiatives.

Limitations and Implications for Research

This study has several limitations. The data were collected through an open-ended online questionnaire, which limited in-depth exploration compared to interviews or classroom observations. The findings are also context-specific to Indonesian primary teachers and should be interpreted cautiously in other educational contexts. Future research could connect these findings with ESD competency frameworks and explore classroom implementation through interviews, observations, or intervention studies. Involving other participants, such as students, may also deepen understanding of how water sustainability is practiced in primary education.

Acknowledgments

The authors would like to express their gratitude to Universitas Negeri Semarang (UNNES) for the financial support provided for this research. The authors also thank all participating primary school teachers for their valuable time and insights, as well as the institutions and professional teacher networks that supported the distribution of the questionnaire.

Funding

This research was funded by Universitas Negeri Semarang (UNNES)

Ethics Statements

This study was conducted in accordance with ethical research standards. Participation was voluntary, and informed consent was obtained from all participants prior to data collection. Respondents were assured of anonymity and confidentiality.

Conflict of Interest

The authors declare no conflict of interest.

Informed Consent

Informed consent was obtained from all participants prior to data collection.

Data availability

The datasets generated and/or analyzed during the current study are available from the corresponding author upon reasonable request.

AI Disclosure Statement

The authors used artificial intelligence to assist with language refinement and improve the clarity and readability of the manuscript. All interpretations, analyses, and conclusions presented in this article are solely those of the authors, who take full responsibility for the content of the manuscript.

References

- Alsina, Á., & Vázquez, C. (2025). Professional development and teacher agency in mathematics teacher education for sustainability. *Mathematics Education Research Journal*, 37(2), 215–238. <https://doi.org/10.1007/s13394-024-00488-y>
- Ampartzaki, M., Kalogiannakis, M., & Papadakis, S. (2021). Deepening our knowledge about sustainability education in the early years: Lessons from a water project. *Education Sciences*, 11(6), 251. <https://doi.org/10.3390/educsci11060251>

- Annamalai, N., Zakarneh, B., Kabilan, M. K., & Karmina, S. (2025). An integration of the sustainable development goals (SDGs) in educational textbooks: The case of the Malaysian high school English textbooks. *World Journal of English Language*, 15(8), 137–148. <https://doi.org/10.5430/wjel.v15n8p137>
- Blom, R., & Karrow, D. D. (2024). Environmental and sustainability education in teacher education research: An international scoping review of the literature. *International Journal of Sustainability in Higher Education*, 25(1), 903–926. <https://doi.org/10.1108/IJSHE-07-2023-0288>
- Boon, H. J. (2024). A comprehensive approach to water literacy in the context of climate change. *Education Sciences*, 14, 564. <https://doi.org/10.3390/educsci14060564>
- BPS-Statistics Indonesia. (2025). *Environment statistics of Indonesia 2025*. BPS-Statistics Indonesia.
- Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77–101. <https://doi.org/10.1191/1478088706qp063oa>
- Braun, V., Clarke, V., & Hayfield, N. (2022). 'A starting point for your journey, not a map': Nikki Hayfield in conversation with Virginia Braun and Victoria Clarke about thematic analysis. *Qualitative Research in Psychology*, 19(2), 424–445. <https://doi.org/10.1080/14780887.2019.1670765>
- Cabarcas-Fuentes, A., & Rodríguez-Chueca, J. (2026). Making it crystal clear: Educating for water sustainability through service-learning. *Education for Chemical Engineers*, 54(1), 100499. <https://doi.org/10.1016/j.ece.2026.100499>
- Calikusu, S. & Sabanci, O. (2025). Students' perspectives on water literacy: The case of 8th grade students. *Journal of Pedagogical Research*, 9(5), 139–158. <https://doi.org/10.33902/JPR.2025358128>
- Do, N. T., & Thinh, M. P. (2025). Integrating sustainable development goals into high school curricula: teachers' perceptions and challenges in the Mekong Delta, Vietnam. *International Research in Geographical and Environmental Education*, 0(0), 1–17. <https://doi.org/10.1080/10382046.2025.2485917>
- Echegoyen, Y., & Martín, A. (2021). A holistic approach to education for sustainability: Ecofeminism as a tool to enhance sustainability attitudes in pre-service teachers. *Journal of Teacher Education for Sustainability*, 23(1), 5–21. <https://doi.org/10.2478/jtes-2021-0002>
- Elliott, C. B. (1996). Pedagogical reasoning: Understanding teacher decision making in a cognitive apprenticeship setting. , *Literacy, Teaching and Learning*, 2(2), 75–91.
- Espinoza-Ramos, G., & Capucci-Polzin, A. (2025). Embedding Sustainability in curricula: A case study of an education for sustainable development (ESD) framework at the westminster business school. In *CSR, Sustainability, Ethics and Governance: Vol. Part F325* (pp. 279–299). https://doi.org/10.1007/978-3-031-86337-0_16
- Ferrer-Est, M., & Chalmeta, R. (2021). Integrating sustainable development goals in educational institutions. *The International Journal of Management Education*, 19(February). <https://doi.org/10.1016/j.ijme.2021.100494>
- Fikri, R. A., Suwono, H., Susilo, H., & Kusairi, S. (2026). Fostering students ' creative thinking through inquiry - design-based STEM water purification project. *Journal of Pedagogical Research, Advance online*. <https://doi.org/10.33902/JPR.202638877>
- Forsler, A., Nilsson, P., & Walan, S. (2024). Capturing and developing teachers' pedagogical content knowledge in sustainable development using content representation and video - based reflection. *Research in Science Education*, 54(3), 393–412. <https://doi.org/10.1007/s11165-023-10149-y>
- Forsler, A., Nilsson, P., & Walan, S. (2025). Teaching for sustainable development in vocational education. *Journal of Biological Education*, 59(2), 268–281. <https://doi.org/10.1080/00219266.2024.2320110>
- Garcia, X., Domene, E., Goenaga, X., Benitez, A. R., Satorras, M., Acuña, V., Ruiz, A. M., Boada, I., & Corominas, L. (2025). Evaluating the effectiveness of a serious game to educate children about the urban water cycle. *British Journal of Educational Technology*, 2023, 1–20. <https://doi.org/10.1111/bjet.13578>
- Green, M., & Somerville, M. (2015). Sustainability education: Researching practice in primary schools. *Environmental Education Research*, 21(6), 832–845. <https://doi.org/10.1080/13504622.2014.923382>
- Hamwy, N., Bruder, J., Sellami, A., & Romanowski, M. H. (2023). Challenges to teachers implementing sustainable development goals frameworks in Qatar. *Sustainability (Switzerland)*, 15(15), 1–19. <https://doi.org/10.3390/su15151479>

- Haque, S. E., Snigdho, F. S., & Tasneem, N. (2025). *Chapter 11 - Water sustainability: a review of advances in water quality management technologies* (A. K. Dubey, A. L. Srivastav, A. Kumar, F. P. Garcia Marquez, & D. A. B. T.-C. A. for W. S. Giannakoudakis (eds.); pp. 195–214). Elsevier. <https://doi.org/10.1016/B978-0-443-33321-7.00018-4>
- Hashim, A. (2023). Literacy and lifelong learning in the twenty-first century: Development of multilingualism and multiliteracies in ASEAN. In *International Handbook on Education Development in the Asia-Pacific* (pp. 493–507). https://doi.org/10.1007/978-981-19-6887-7_36
- Imaduddin, M., & Eilks, I. (2024). A scoping review and bibliometric analysis of educational research on water literacy and water education. *Sustainable Chemistry and Pharmacy*, 42(3), 101833. <https://doi.org/10.1016/j.scp.2024.101833>
- Inoue, M., O’Gorman, L., & Davis, J. (2016). Investigating early childhood teachers’ understandings of and practices in education for sustainability in Queensland: A Japan-Australia research collaboration. *Australian Journal of Environmental Education*, 32(2), 174–191. <https://doi.org/10.1017/ae.2016.4>
- Irimia, O., Tomozei, C., Panainte-Lehadus, M., Chitimus, D., Nedeff, F., Barsan, N., Mosnegutu, E., & Mirila, D. (2025). Understanding primary school students’ knowledge and attitudes towards water management: Insights from environmental education. *Societies*, 15(4). <https://doi.org/10.3390/soc15040109>
- Kahn, R. (2010). Critical pedagogy, ecoliteracy and planetary crisis: The ecopedagogy movement: A Book Review Symposium. *Journal for Critical Education Policy Studies*, 8(1).
- Kwee, C. T. T., & Dos Santos, L. M. (2023). An international study of high school teachers’ experience of incorporating water resources in their teaching. *Frontiers in Education*, 7, 1065228. <https://doi.org/10.3389/educ.2022.1065228>
- Mackiewicz, J. (2018). *Writing Center Talk over Time a Mixed-Method Study*. Routledge.
- McCarroll, M., & Hamann, H. (2020). What we know about water: A water literacy review. *Water (Switzerland)*, 12(10). <https://doi.org/10.3390/w12102803>
- Misiaszek, G. W. (2015). Ecopedagogy and citizenship in the age of globalisation: Connections between environmental and global citizenship education to save the planet. *European Journal of Education*. <https://doi.org/10.1111/ejed.12138>
- Mostacedo-Marasovic, S.-J., Mott, B. C., White, H., & Forbes, C. T. (2022). Towards water literacy: an interdisciplinary analysis of standards for teaching and learning about humans and Water. *Disciplinary and Interdisciplinary Science Education Research*, 4(1). <https://doi.org/10.1186/s43031-022-00065-y>
- Pozo-Muñoz, M. P., Martín-Gámez, C., & Velasco-Martínez, L. C. (2023). Analysis and assessment of the environmental awareness of primary school pupils on the management and sustainable use of water. *International Research in Geographical and Environmental Education*, 32(3), 217–233. <https://doi.org/10.1080/10382046.2022.2154963>
- Praveena, S. M., & Themudu, S. (2022). Exploring Water Conservation Awareness Level Among Primary School Children From Melaka (Malaysia). *Malaysian Journal of Medicine and Health Sciences*, 18, 75–81.
- Ramírez-Segado, A., Castro-velásquez, F. E., & Benarroch, A. (2025). *Water Sustainability among Secondary Education Teachers: Perspectives from Two Hispanic Cities*. 1–16. <https://doi.org/10.1017/ae.2025.10084>
- Ramírez-Segado, A., Rodríguez-serrano, M., & Benarroch, A. B. (2021). El agua en la literatura educativa de las dos últimas décadas . Una revisión sistemática. *Revista Eureka Sobre Enseñanza y Divulgación de Las Ciencias*, 18(1), 1107. https://doi.org/10.25267/Rev_Eureka_ensen_divulg_cienc.2021.v18.i1.1107
- Rey, R. Del, Ojeda, M., Mora-merchán, J. A., Nieves, M., Morgado, B., & Lasaga, M. J. (2022). Environmental education : Effects on knowledge , attitudes and perceptions , and gender differences. *International Research in Geographical and Environmental Education*, 31(4), 282–303. <https://doi.org/10.1080/10382046.2021.1977004>
- Romulo, C., Venkataraman, B., Caplow, S., Ajgaonkar, S., Allen, C. R., Anandhi, A., Anderson, S. W., Azzarello, C. B., Brundiers, K., Blavascunas, E., Dauer, J. M., Druckenbrod, D. L., Fairchild, E., & Horne, L. R. (2024). Implementing interdisciplinary sustainability. *Humanities and Social Sciences Communications*, 1–17. <https://doi.org/10.1057/s41599-024-03332-7>

- Rosmiati, Satriawan, M., Fanny, A., Satianingsih, R., Utari, S., Ningsih, E., Rasyiida, G., Anandha, A., & Anam, P. (2025). Design of reflective-sustainable development education oriented project-based learning platform for primary school students in Indonesia. *Multidisciplinary Science Journal*, 7(9). <https://doi.org/10.31893/multiscience.2025444>
- Ruiz-Garzón, F., Olmos-Gómez, M. D. C., & Estrada-Vidal, L. I. (2021). Perceptions of teachers in training on water issues and their relationship to the SDGs. *Sustainability (Switzerland)*, 13(9). <https://doi.org/10.3390/su13095043>
- Scarborough, D. L. A., Hall, R. D., & Vanderkruk, K. E. N. (2022). Laboratory research projects in undergraduate environmental and analytical chemistry. *Journal of Chemical Education*, 99(4), 1672–1681. <https://doi.org/10.1021/acs.jchemed.1c01181>
- Shulman, L. S. (1987). Knowledge and teaching: Foundations of new reform. *Harvard Educational Review*, 57(1).
- Siraj, J., & Pramling, I. (2016). Education for sustainable development in early childhood care and education: An introduction. *International Perspectives on Early Childhood Education and Development*, 14, 1–15. https://doi.org/10.1007/978-3-319-42208-4_1
- Stephenson Reaves, J. (2023). Connecting local questions to global issues: An investigation with primary pre-service teachers BT. In G. A. Buck, V. Dimitrieska, & V. L. Akerson (Eds.), *Internationalizing rural Science Teacher Preparation: Action Research for Global Competency* (pp.131–145). Springer International Publishing. https://doi.org/10.1007/978-3-031-46073-9_9
- The Minister of Environment and Forestry, R. of I. (2023). *Statistik 2023 Kementerian Lingkungan Hidup dan Kehutanan: Vol. IX* (Edi Sulisty H.Susetyo (ed.). Pusat Data dan Informasi Kementerian Lingkungan Hidup dan Kehutanan.
- Tsiaras, S., & Chavenetidou, M. (2024). Views and attitudes of primary school students regarding water conservation in the municipality of Thessaloniki, Greece. *E3S Web of Conferences*, 585, 3–6. <https://doi.org/10.1051/e3sconf/202458510005>
- UN-Water. (2016). *Water and sanitation interlinkages across the 2030 agenda for sustainable development* (W. J. Rogers (ed.)). UN-Water.
- UNESCO. (2020). United Nations World Water Development Report 2020: Water and climate change. UNESCO. www.unesco.org/open-access/termsuse-ccbysa-en
- Uygun, A.G. & Bayrakci, M. (2025). Teachers' perspectives on environmental education practices in schools from the perspective of sustainable schools. *Political Economy and Management of Education*, 6(1), 18–44. <https://dergipark.org.tr/en/pub/peme>
- Valenzuela-Morales, G. Y., Hernández-Téllez, M., Ruiz-Gómez, M. L., Gómez-Albores, M. A., Arévalo-Mejía, R., & Mastachi-Loza, C. A. (2022). Water conservation education in primary schools: The case of the Nenetzingo River catchment, Mexico. *Sustainability (Switzerland)*, 14(4). <https://doi.org/10.3390/su14042402>
- Vesterinen, M. (2024). Sustainability competences in environmental education: Research on guidebooks for teachers at Finnish primary schools. *Cogent Education*, 11(1). <https://doi.org/10.1080/2331186X.2023.2286120>
- Višnjic-Jevtic, A., Lepicnik Vodopivec, J., Pribisev Beleslin, T., & Sindic, A. (2022). Unmasking sustainability in early childhood education: Teachers' voices from Bosnia and Herzegovina, Croatia, and Slovenia. *International Journal of Early Childhood*, 54(1), 119–137. <https://doi.org/10.1007/s13158-022-00321-2>
- Wang, X., Liu, Z., & Zhang, Y. (2024). Differences in water-saving behaviors among college students: Research based on the theory of planned behavior. *Sustainability (Switzerland)*, 16(23). <https://doi.org/10.3390/su162310182>
- Yang, H. (2021). Application of project based learning in an environmental engineering programme. *International Research Symposium on PBL*, 195–205. <https://www.scopus.com/inward/record.uri?eid=2-s2.0-85113171508&partnerID=40&md5=0b27b4f43899558bd3f3bb56ad33d2f6>
- Zhan, Y., He, R., Wing, W., & So, M. (2019). Developing primary school children's water conversation action competence: A case study in China. *International Journal of Early Years Education*, 27(3), 287–305. <https://doi.org/10.1080/09669760.2018.1548346>

Appendix A.*Examples of the coding and theme development process***Table A1.***Examples of the coding and theme development for teachers' rationales*

Participant excerpt	Initial code	Theme
"I think children should learn about water conservation from an early age so they understand how important water is for life."	Early water awareness	Raising environmental awareness
"If students learn how to save water when they are young, it can become a habit that continues into adulthood."	Eco-conscious generation	Raising environmental awareness
"Children should understand that clean water is important and needs to be protected so it remains available in the future."	Availability of clean water	Alignment with SDGs
"Future generations will also need clean water, so students should learn why it needs to be conserved."	Water sustainability	Alignment with SDGs
"Besides learning about water, students can also learn to appreciate and be grateful for the water they have."	Good and being grateful	Character and values education
"Students should be encouraged not to waste water and to use it only when they really need it."	Use water wisely	Character and values education
"I think this topic can help children develop good habits that they can practice every day."	Build positive character	Character and values education
"Children should learn that protecting water resources is everyone's responsibility."	Environmental stewardship	Character and values education
"Students use water every day when they take a bath, so it is easy for them to relate to this topic."	Bathing	Relevance to daily life
"Drinking water is part of students' everyday lives, which makes the topic meaningful for them."	Drinking	Relevance to daily life
"Students use water when washing their hands, clothes, and other things at home."	Washing	Relevance to daily life
"Many students help water plants at home or at school, so they already have experience related to water use."	Watering plants	Relevance to daily life
"Without water, people, animals, and plants would not be able to survive."	Water is a source of life	Supporting science literacy
"Students need to know that water is important for keeping nature and ecosystems balanced."	Importance of ecosystem balance	Supporting science literacy

Table A2.*Examples of the coding and theme development for teachers' perceived opportunities*

Participant excerpt	Initial code	Theme
"When discussing floods and droughts, teachers can also introduce the importance of water conservation."	Natural events	Integration into primary science topics
"Water sustainability can be linked to lessons about a clean and healthy environment."	Clean and healthy environment	Integration into primary science topics
"The topic fits well with lessons about ecosystems and the water cycle."	Ecosystem	Integration into primary science topics
"When learning about natural resources, students can also discuss how water should be used and protected."	Natural resources and environment	Integration into primary science topics
"Water can be discussed as part of maintaining ecosystem balance."	Harmony in the ecosystem	Integration into primary science topics
"Environmental problems affecting the Earth can be connected to water sustainability issues."	Our Earth is in danger	Integration into primary science topics
"Students can discuss their responsibilities in protecting natural resources, including water."	Citizens' responsibility in protecting natural resources	Cross-subject integration through environmental themes
"Students can write stories, poems, or short texts about the importance of saving water."	Write an appeal text, create a poem, or a story	Cross-subject integration through environmental themes
"Daily water use can be used as an example when learning measurement and volume."	Calculating volume	Cross-subject integration through environmental themes
"Students can create posters or drawings that promote water conservation."	Creating artwork illustrating water conservation messages	Cross-subject integration through environmental themes
"Water shortages can help students understand how environmental problems affect people's lives."	Sustainable development and environmental change	Cross-subject integration through environmental themes
"Students can organize simple campaigns to encourage water-saving habits."	Water-saving campaign project	Integration through thematic projects
"Students can record how much water they use at home during the week."	Observation of water use in daily life	Integration through thematic projects
"School clean-up activities can be linked to protecting local water resources."	Environmental cleanliness project	Integration through thematic projects

Table A3.
Examples of the coding and theme development for teachers' perceived barriers

Participant excerpt	Initial code	Theme
"Our school does not have enough facilities to support environmental observation activities."	Lack of supporting facilities at school	Limited resources and facilities
"There are still limited teaching materials and media related to water sustainability."	Limited learning resources	Limited resources and facilities
"Not all teachers are familiar with water sustainability concepts."	Lack of understanding of water sustainability	Limited teacher knowledge and confidence
"It can be challenging to explain water sustainability in language that young children can easily understand."	Difficulty explaining sustainability concepts	Limited teacher knowledge and confidence
"I am not always confident discussing environmental issues beyond the textbook."	Low confidence in teaching sustainability	Limited teacher knowledge and confidence
"There is already a lot of content that needs to be covered in the curriculum."	Dense curriculum focus	Curriculum constraints
"Teachers often prioritize topics that are directly assessed in school examinations."	Competing curriculum priorities	Curriculum constraints
"Students who always have easy access to water often do not see why conservation is important."	Perceived abundance of water	Environmental context
"Students from different environments often have different levels of awareness about water issues."	Local environmental influences	Environmental context
"Children who have never experienced water shortages may find the topic less relevant."	Limited experience with water-related problems	Environmental context
"Project activities usually require more time than regular classroom lessons."	Limited learning time at school	Time limitation
"The available lesson time is often not enough for practical activities."	Insufficient instructional time	Time limitation
"It can be challenging to find teaching methods that are both engaging and suitable for young learners."	Difficulty selecting appropriate methods	Method uses
"Teachers need practical and engaging ways to introduce water sustainability to young learners."	Difficulty packaging explanations	Method uses
"Some learning approaches work well for older students but are difficult to adapt for primary school children."	Challenges in adapting instructional strategies	Method uses

Note. The excerpts presented in this appendix are illustrative examples selected from participants' responses to demonstrate the coding and theme development process. The excerpts were translated from Indonesian into English by the authors. Coding was conducted inductively and iteratively, and themes were refined through repeated comparison and review using NVivo.

A Case Study of Kindergarten Teachers' Science Practices: Reflections of Knowledge and Attitude*

Nefise Özok Bulut^{a,*}, Meral Ören^b

Received : 2 April 2026
Revised : 29 May 2026
Accepted : 5 June 2026
DOI : 10.26822/iejee.2026.463

^{a,*} **Corresponding Author:** Nefise Özok Bulut, Assistant Professor, Department of Child Development, İstanbul Gedik University, İstanbul, Türkiye.
E-mail: nefise.ozok@gedik.edu.tr
ORCID: <https://orcid.org/0000-0002-2070-7797>

^b Meral Ören, Assistant Professor, Department of Early Childhood Education, Anadolu University, Eskişehir, Türkiye.
E-mail: moren@anadolu.edu.tr
ORCID: <https://orcid.org/0000-0003-1583-6874>

* This study was derived from the first author's master's thesis completed at the Department of Elementary Education, Graduate School, Anadolu University, Eskişehir, Türkiye, and was previously presented as an oral presentation at the International Education Congress held on November 17-19, 2022.

Abstract

It is becoming increasingly recognized and accepted that providing children with quality science experiences during their early years can positively impact their science achievement and process skills in their later years. To provide such experiences, it is critical to explore how early childhood teachers' attitudes toward science and their level of knowledge in science impact their science teaching in the classroom. The study aimed to examine the science practices of kindergarten teachers with different levels of scientific knowledge and attitudes toward teaching science employing a holistic multiple-case study. The participants were two kindergarten teachers with different levels of scientific knowledge and attitudes toward teaching science. They were selected through an extreme case sampling method using the Early Childhood Teachers' Attitudes Toward Science Teaching Scale and the Achievement Test on Scientific Concepts results. The data were collected through interviews and observation; and analyzed through deductive thematic analysis. The results indicated that science knowledge and attitudes of teachers were related to; the frequency of science activities, the resources they used, planning processes, teaching methods, making accurate scientific explanations, and child involvement in science activities.

Keywords:

Early Childhood Education, Kindergarten Teachers, Science Education, Scientific Knowledge, Teachers' Attitudes Towards Teaching Science

Introduction

According to the results of the Programme for International Student Assessment (PISA) tests, Türkiye's performance in science literacy, though improving, continues to lag statistically behind the OECD average (OECD, 2019, 2022). This persistent gap underscores the critical need to enhance science education across all developmental stages, beginning with early childhood. Research findings, particularly from the Trends in International Mathematics and Science Study (TIMSS), affirm a positive correlation between the duration of preschool education and later science literacy outcomes (MoNE, 2016). Furthermore, quality, enriched science experiences provided in preschool are shown to be foundational for children's future scientific achievement and the development of essential scientific



www.iejee.com
ISSN: 1307-9298

2026 Published by KURA Education & Publishing.
This is an open access article under the CC BY-NC-ND license. (<https://creativecommons.org/licenses/by/4.0/>)

process skills (Genç-Kumtepe et al., 2009; Morgan et al., 2016).

Effective science instruction hinges upon the quality of the kindergarten teacher. While studies conducted in Türkiye consistently report that early childhood teachers generally hold positive attitudes toward teaching science, this enthusiasm often coexists with documented deficiencies in their science content knowledge (CK) and low professional self-efficacy (Mercan & Kutluca, 2021; Olgan, 2015; Polat et al., 2021). This disjuncture presents a primary challenge in the field: Although positive attitudes can be linked to a higher frequency of implementing science activities (Genç-Kumtepe et al., 2013), insufficient scientific knowledge often creates significant difficulties in executing these activities effectively and authentically (Akyol & Birinci-Konur, 2018; Babaroğlu & Okur-Metwalley, 2018). Crucially, the underlying reasons why teachers avoid or oversimplify science often extend beyond structural limitations to affective and self-efficacy-related barriers. Research suggests that early childhood educators frequently experience discomfort, anxiety, and a lack of confidence when teaching science, particularly when they perceive their science content knowledge as insufficient (Roberts & Collins, 2026; Simsar & Davidson, 2020). Such feelings of inadequacy may lead teachers to avoid science activities, postpone scientific investigations, or transform inquiry-based experiences into highly structured and predictable routines, thereby limiting children's opportunities for authentic scientific exploration (O'Connor et al., 2021; Roberts & Collins, 2026).

Much of the existing research primarily emphasizes that professional development should prioritize raising teachers' knowledge and attitudes, positing these factors as the necessary drivers for better science practice frequency and quality (Barenthien et al., 2020; Loxley et al., 2016; Oppermann et al., 2021; Pendergast et al., 2017). However, the complexity of this relationship is often overlooked. Training programs in Türkiye currently offer only one science education course, leading teachers to perceive their undergraduate preparation as insufficient (Mercan & Kutluca, 2021). Studies show that exposure to more methods courses or professional development is linked to greater proficiency and frequency of science integration (Oppermann et al., 2021; Sağkes, 2014). Crucially, given that early childhood teachers can inadvertently transmit scientific misconceptions, they must possess both content knowledge (CK) and pedagogical knowledge (PK) for science (Gerde et al., 2018; Greenfield et al., 2009; Kallery & Psillos, 2001; McCray & Chen, 2012).

Grounded in Gudmundsdottir and Shulman's (1987) PCK model, this study examines how early

childhood teachers transform subject matter into teachable content. Within this model, PCK functions as an integrated framework encompassing core components: deep content knowledge, specialized pedagogical techniques, and learner-tailored instructional strategies that accommodate children's developmental readiness and preconceptions (Nilsson & Loughran, 2012; Park & Oliver, 2008). In early childhood science, these components dictate how a teacher moves beyond generic teaching methods to deploy domain-specific scaffolding. Scientific knowledge and attitudes serve as critical precursors to this process; deficits in either inevitably constrain PCK, replacing genuine inquiry with superficial, recipe-following activities. Consequently, this research investigates not merely the presence of these elements, but how they interact to shape enacted PCK in the classroom. Crucially, because a teacher's enacted PCK directly delineates the boundaries of the classroom learning environment, it fundamentally shapes child outcomes—determining whether children exhibit active inquiry and high agency or succumb to passive disengagement and reinforced misconceptions during science experiences.

Furthermore, these classroom enactments do not occur in a vacuum but are highly influenced by the shifting curriculum and socio-cultural environment. While the recent curriculum revisions by the Ministry of National Education (MoNE) aim to modernize early childhood practices (MoNE, 2024), contemporary literature highlights that the post-COVID-19 digital transformation and the introduction of hybrid learning approaches have fundamentally reshaped early childhood science landscapes in Türkiye, creating additional demands for techno-pedagogical PCK readiness (Aktaş, 2024; Öztürk, 2025). Concurrently, the unique socio-cultural dynamics of school-family relationships in the Turkish context introduce a distinct mechanism of parental delegation or pedagogical compensation. In many local communities, parents heavily delegate scientific and academic accountability to the school, expecting early childhood teachers to act as the primary or sole source of formal knowledge acquisition. This cultural adaptation places an immense implicit pressure on teachers' enacted PCK, making their perceived scientific competence and classroom practices a critical junction for both institutional and familial expectations (Snape & Fox, 2021), particularly within localized private and institutional early childhood sectors where consumerist parental demands are highly pronounced.

While prior quantitative studies consistently map the correlation between teacher characteristics (CK, self-efficacy, and attitudes) and the frequency of science instruction (Mercan & Kutluca, 2021; Olgan, 2015; Polat et al., 2021), there is a profound critical gap

in understanding the qualitative, contextual dynamics of this relationship. The field currently lacks in-depth, systematic case studies that move beyond self-reports to analyze how knowledge and attitudinal deficiencies are reflected in the nuanced processes of planning, pedagogical choices, and real-time interaction with children during science lessons.

To address this critical gap, the purpose of this qualitative comparative case study is to examine, in-depth, the science planning and implementation practices of two kindergarten teachers possessing distinctly different profiles of science content knowledge and professional self-efficacy, while concurrently exploring how these disparate pedagogical environments shape kindergarteners' classroom reactions. The results are expected to contribute significantly to the refinement of pre-service teacher training programs, guide context-specific in-service training efforts, and inform curriculum development studies by providing rich data. To achieve this objective, the study addresses the following research questions:

1. How do kindergarten teachers with different levels of science knowledge and attitudes towards science education:
 - a. Plan science activities?
 - b. Implement science activities?
2. How do kindergarteners react during the science practices of kindergarten teachers?

Method

Research Design

The research design for this study employed a holistic multiple-case study design. This approach was selected to provide an intensive, contextual examination of a contemporary phenomenon (the relationship between teacher characteristics and practice) within its real-life setting (Creswell & Creswell, 2018). The decision to utilize a multiple-case study was methodologically driven by the need to explore variations and similarities in science teaching practices across distinctly different profiles of teacher knowledge and attitude.

In accordance with Yin's (2018) framework, a holistic design was applied to analyze each teacher-school context as a single, unified case, explicitly examining how individual subject matter boundaries, pedagogical content knowledge (PCK), and parental interactions intersect in daily instruction. Because the first author operated as a participant observer over a 16-week period, potential researcher influence and interpretive bias were meticulously managed through rigorous reflexivity. The researcher maintained a detailed diary of analytical memos to intentionally bracket a priori professional assumptions, ensuring

that findings remained closely grounded in the participants' authentic experiences rather than the researcher's expectations (Gray, 2004). Consequently, the resulting qualitative conclusions are presented non-definitively and with scientific caution, acknowledging that these bounded cases illuminate complex pedagogical dynamics rather than offering universal generalizations.

Participants

To select the cases, 119 volunteering teachers from schools in three districts of Istanbul, one of the major cities in Türkiye, completed the Early Childhood Teachers' Attitudes Toward Science Teaching Scale and the Achievement Test on Scientific Concepts. Using the scale scores, one teacher from both ends was selected through extreme case sampling using purposeful sampling methods. An extreme case sampling strategy (Silverman, 2015), a form of purposeful sampling, was employed to maximize the variation in teacher practices and enable the analysis of contrasting cases. To guide the reader, the teacher with the highest scores in the study is given the pseudonym 'Hale', starting with the initials of the word 'high', whereas the teacher with low scores is given the pseudonym 'Leyla', which begins with the initial letter of the word 'low'. All participants were provided with a participant information statement and informed consent.

Leyla was strategically selected to represent a profile slightly below the group average, deliberately avoiding extreme low-scoring outliers whose practices may be idiosyncratic. By focusing on this 'moderate but insufficient' profile, the study captures the reality of a substantial segment of the teacher population, thereby enhancing the contextual transferability of the findings regarding typical challenges. Table 1 presents Hale and Leyla's scores relative to the larger cohort ($N = 119$).

Table 1.
Participants' achievement test and attitude scale scores

	Hale	Leyla	Minimum	Maximum	Average
Attitude scale score	62	47	29	65	47.67
Achievement test score	19	7	7	19	13.47

Hale received the highest scores on both scales, while Leyla scored well below average on the Achievement Test on Scientific Concepts and had an average score on the Early Childhood Teachers' Attitudes Toward Science Teaching Scale. The characteristics of the teachers are demonstrated in Table 2.

Table 2.*Characteristics of the participants*

	Hale	Leyla
Gender	Female	Female
Institution	Kindergarten	Kindergarten
Number of children in the class	23	23
Age	38	36
Professional experience	6	15
Undergraduate program	Early childhood education	Early childhood education

Both teachers operated within public primary schools adhering to the national MoNE (Ministry of National Education) curriculum. Kindergarten in this paper refers to the final year of formal schooling before first grade, as commonly used in the USA context, serving children aged 57-68 months at enrollment. Both schools were economically located in lower-middle-class neighborhoods and had a class size of approximately 25-28 students.

Schools with similar physical characteristics and resources were selected as they could influence science practices. However, some differences exist in science practices. Hale conducts formal science activities three to four times a week, often utilizing hands-on materials. Leyla conducts science activities two to three times a week, and her practices were observed to rely more on readily available classroom materials rather than specialized science resources.

Data Collection and Procedures

Two instruments guided the selection of case study participants. First, the Early Childhood Teachers' Attitudes towards Science Teaching Scale (Cho et al., 2003; adapted by Ünal et al., 2010) was used to assess self-development and self-efficacy. Reliability coefficients were high for both the original adaptation ($\alpha = 0.84$) and the current study ($\alpha = 0.89$). Second, the Achievement Test on Scientific Concepts (Güney, 2016) measured content knowledge. While the original test demonstrated a reliability of $\alpha = 0.77$, the current study yielded a coefficient of $\alpha = 0.67$.

Following participant selection, data were collected through 16 weeks of classroom observations and interviews (November 2019 – February 2020). Weekly semi-structured micro-interviews (Merriam, 2018) were conducted immediately before and after observations. The interview questions and prompts were developed based on a comprehensive literature review focusing on teacher self-efficacy (Bandura, 1997) and PCK dimensions (Gudmundsdottir & Shulman, 1987). To ensure content validity and pedagogical appropriateness, the initial draft of the questions was evaluated by three independent experts holding

doctoral degrees in early childhood education. Following their feedback, minor adjustments were made to the phrasing of the questions to enhance clarity. A pilot application was subsequently conducted with one non-participating preschool teacher to test the flow, comprehensibility, and duration of the micro-interviews, which confirmed the instrument's readiness for the field. These brief sessions (<10 mins) captured in-the-moment planning rationales and immediate reflections on practice (e.g., "Which science topics cause you anxiety?" or "How do you handle unanswerable questions from children?"). The first author utilized a participant observer technique, engaging in free play but remaining an observer during science activities. The focus extended beyond the teacher to include the verbal and nonverbal behaviors of interacting children.

Due to school administration policies, institutional research ethics, and strict child privacy regulations within early childhood settings in Türkiye, video recording of the children and classroom environments was not permitted. Consequently, data triangulation was rigorously implemented to ensure trustworthiness. Audio recordings of the interviews were transcribed verbatim and supplemented by detailed, chronological field notes capturing non-verbal cues (e.g., body language, attention shifts) observed during the classroom sessions. Additionally, a researcher's diary was maintained to record analytical memos and facilitate reflexivity (Gray, 2004). This process involved bracketing a priori beliefs to ensure interpretations remained grounded in the participants' perspectives rather than the researcher's assumptions.

Data Analysis

Qualitative data were analyzed using thematic analysis following the six-phase framework outlined by Braun and Clarke (2006). To balance theoretical coherence with empirical sensitivity, a hybrid coding strategy was employed:

Major themes were established a priori based on the study's research questions and the PCK framework (e.g., planning process, teaching process). This structure anchored the findings within the established theoretical scope.

Sub-themes and specific codes emerged organically from the data. Through deep immersion in verbatim transcripts and observation notes, codes such as difficulties encountered or children's reactions were derived directly from participants' experiences, capturing nuances not explicitly dictated by the PCK model. This dual approach ensured that the analysis remained theoretically grounded while being open to context-specific phenomena.

Validity and Reliability

To maximize the trustworthiness of this qualitative research, we implemented multiple methodological procedures (Creswell & Creswell, 2018; Merriam, 2018), addressing all four key criteria:

Established through data triangulation (cross-referencing interviews, observation notes, and the researcher's journal) and long-term engagement (16 observations per case). Furthermore, member checking was conducted by sending detailed summaries to Hale and Leyla for review and confirmation of accuracy.

Ensured via inter-coder reliability. A second, independent coder analyzed 10% of the total transcriptions. The agreement rate was calculated using the formula by Miles and Huberman (1994) across 75 total coding instances, yielding a final reliability score of 92%. Discrepancies were resolved through discussion to reach a consensus.

Supported by providing thick descriptions of the teachers' practices and the detailed account of the research context. Enhanced through researcher reflexivity, involving the systematic examination of the researcher's a priori beliefs and positionality, as maintained in the researcher's diary.

To further ensure transparency, the first author's positionality and potential observer effects must be critically explicated. As a doctoral researcher well-versed in early childhood science education, the first author entered the field with an a priori assumption that high-quality, inquiry-based science is critical for young children, which initially created a subjective expectation toward witnessing inquiry-driven practices. Acknowledging this bias, the researcher rigorously utilized analytical memos to bracket these personal expectations during data analysis. Furthermore, the prolonged 16-week presence of the researcher inevitably introduced subtle observer effects on classroom dynamics.

Most notably, the teacher with limited science expertise explicitly increased the frequency of her science activities during the initial weeks of the observation period, demonstrating a clear reaction to being observed. Rather than discarding this reactivity, the researcher treated this behavioral shift as an empirical finding in itself, documenting how the teacher initially relied on structured, internet-sourced demonstration kits to artificially meet the perceived professional expectations of the ongoing study. Over the extended timeline, however, this acute observer effect gradually diminished; the prolonged presence allowed the researcher to become a familiar, non-threatening element of the classroom environment. This long-term immersion eventually prompted both

teachers to revert to their authentic pedagogical routines, exposing the deep-seated pedagogical gaps, anxieties, and compensatory behaviors (such as relying on the 'parent of the week') that form the core of this study's findings. By critically reporting these reflexive shifts, the study embraces researcher-participant dynamics as a transparent dimension of qualitative rigor.

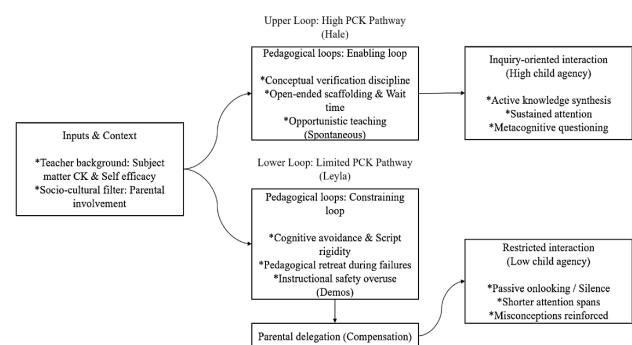
Ethical Procedures

This study was approved by the Anadolu University Social and Human Sciences Scientific Research and Publication Ethics Committee (Date: March 28, 2019, Protocol Number: 22483). Institutional permissions were secured from the participating private preschool administrations prior to data collection. Written informed consent was obtained from both participating teachers, and written parental consent alongside verbal child assent was secured for all participating children to ensure strict compliance with ethical standards and absolute confidentiality.

Results

As a result of the analysis obtained through micro-interviews, systematic observations, and the researcher's diary, the findings were inductively grouped under two main themes and eight sub-themes. To make the operational complexity of these qualitative findings immediately digestible for the reader, a comprehensive conceptual framework summarizing the relationship between the teachers' Pedagogical Content Knowledge (PCK) levels and the resulting classroom child interactions is illustrated in Figure 1. This model visually bifurcates the data into two distinct pedagogical pathways based on teacher competence.

Figure 1.
Conceptual framework of PCK pathways and child interaction dynamics



As a result of the analysis obtained through interviews and observations, the findings were grouped under two main themes and eight sub-themes as shown in Table 3.

The data analysis, based on a comprehensive set of micro-interviews, systematic observations, and researcher's diary yielded a rich comparative understanding of the two teachers' science teaching practices. The findings were inductively grouped into two main themes and eight sub-themes, as outlined in Table 3. Furthermore, to immediately provide the reader with a clear structural overview of the key comparisons, the core analytic distinctions between Hale and Leyla are summarized in Table 4, which precedes the thematic presentation.

The following sections detail the findings across the established themes, incorporating direct quotations with source codes (e.g., Observation, Interview, Researcher's diary) to ensure the study's confirmability and analytic rigor.

Table 3.

Themes, sub-themes and codes

Themes	Sub-themes	Codes
Planning process	Resources used	Books, magazines, documentaries, Internet, social media, plans from different sources
	Preparations	Controlling the accuracy of the information, Adapting children's level, Making modifications to ensure active participation of children, Making connections to previous learnings, Preparing interesting materials and learning environments
	Frequency	Two times a week, Three times a week, Four times a week, Unplanned science activities
	Topics	Physical science, earth and space science, life science
	Scientific process skills	Prediction, recall, making associations with background information, cause-and-effect relationships, inference, classification, measurement
Teaching process	Teaching methods and techniques	Question-and-answer, lecture, experiment, demonstration, inquiry
	Difficulties encountered	Attracting children's attention, Making transitions between activities, Classroom management, Adapting information to children's level, Creating a supportive learning environment, Integrating science with other activities
	Children's reactions	Active participation, remaining passive, focusing attention on the activity, answering questions, asking questions, remaining unresponsive, distraction,

Table 4.

Comparative analytical summary of findings

	Hale (Strengths: High Self-Efficacy / Extensive CK)	Leyla (Limitations: Low Self-Efficacy / Limited CK)
Planning and preparation	Conceptual verification discipline: Rigorous checking of information accuracy and use of multiple sources.	Reliance on ready-made materials: Tendency to use pre-made plans with minimal adaptation
Implementation and spontaneity	Opportunistic teaching: Ability to spontaneously convert moments of child interest into conceptual inquiry.	Pedagogical avoidance: Ignoring spontaneous moments that involve uncertainty or move beyond planned content.
Questioning technique	Open-ended inquiry: Providing sufficient wait time and guiding children towards hypothesis generation.	Closed-ended/self-answering: Prioritizing quick closure; answering her own questions; restricting deeper thinking
Questioning technique / Risk management	High conceptual confidence: Unworried about being asked unanswerable questions; actively connecting complex concepts to prior learning	Conceptual gaps: Transmitting inaccurate information due to lack of knowledge and retreating during conceptual failure.
Compensation/Support mechanism	Parental enrichment: Integrating parent-led content into her own inquiry framework.	Parental delegation: Strategically delegating life science activities to parents to reduce personal instructional burden.
Impact on child participation	High agency: Fostering active participation where children exhibit knowledge synthesis and metacognition.	Low agency: Discouraging curiosity and confirming scientific misconceptions through silence or incomplete answers.

Planning Process

Under this theme, the findings related to two sub-themes, namely the resources used by the teachers in the planning process and the preparations they make before the science activities, are given respectively. The findings indicate that differences in teacher competence significantly determined the depth of research and the discipline of conceptual validation.

During the interviews, Hale stated that she explores various sources, including books, magazines, documentaries, the internet, and social media for science activity plan ideas; and mostly plans science activities herself, but sometimes uses the plans from different sources by adapting if necessary. Leyla, on the other hand, stated during the interviews

that she uses magazines, documentaries, and the internet about science, and that she usually uses plans from different sources rather than planning herself. To prepare for the activities, Hale stated that she controls the accuracy of the information, adapts children's level, makes modifications to ensure active participation of children, and makes connections to previous learnings:

"I check the accuracy of the information from a theoretical perspective. I make adaptations according to the age and developmental characteristics of children... I remind them of the knowledge they have already learned. I try to ensure that children participate actively, depending on the activity (Interview)."

Observations indicated that Hale maintained this conceptual verification discipline even when activities were presented by the child of the week. This commitment reflects Hale's high self-efficacy and her professional standard of validating all content introduced into the classroom.

During a "the paper that doesn't get wet" experiment led by a child, Hale did not simply observe. She asked the children, "Do you remember what made the ping-pong ball bounce when we played?" thereby deliberately anchoring the experiment in the previously discussed concept of air pressure (Observation). This proactive intervention exemplifies Hale's ability to instantaneously connect a demonstration to prior learning and a coherent scientific framework.

In contrast, Leyla mentioned preparing interesting materials and learning environments (Interview), but did not specify any attempts to acquire deep conceptual knowledge about the topics she addressed. Researcher's diary entries strongly suggest that this deficiency led to pedagogical retreat when errors occurred:

"Leyla's failure to secure conceptual knowledge increased the risk of conceptual error and diminished her capacity for pedagogical recovery (Researcher's Diary)."

"During a non-exploding balloon experiment, Leyla failed to put enough water in the balloon, causing it to burst. Crucially, she did not discuss the reason for the burst with the children but immediately took a new balloon and repeated the experiment. This demonstrates her avoidance of spontaneously engaging with conceptual failure (Researcher's Diary)."

It was further observed that Leyla did not acquire knowledge regarding the science activities implemented by families, nor did she take an active role. Her presence during a bird-themed activity conducted by a mother was purely observational (Observation), and she admitted in the interview to lacking sufficient knowledge about animal care and characteristics (Interview). This behavior is interpreted

as an early manifestation of her compensatory strategy, where parental delegation is used to manage professional inadequacy.

Teaching Process

Under this theme, findings related to six sub-themes are presented: the frequency of teachers' implementation of science activities, the topics they address, scientific process skills, the teaching methods and techniques they use, the difficulties encountered by teachers, and the reactions of children.

The frequency of implementing science activities

Hale conducts science activities three to four times a week, and in one of these activities, a child does an experiment taught by his/her parents. Hale video records the process while she guides the child, asks questions to the children, and teaches related content knowledge. This practice shows that Hale uses this situation as a source of enrichment.

Leyla conducted science activities two to three times a week. She reported in the interviews that the researcher's presence encouraged her to practice more regularly (Interview). Crucially, Leyla openly stated that she engaged families to perform various science activities "to reduce the subject load and relieve her burden" (Interview). This finding confirms Leyla's use of parental delegation as a core professional compensatory strategy. Each week, this delegation involved the parent of the week conducting a demonstration experiment and teaching about an animal or a plant on a separate day.

A critical distinction emerged in the teachers' ability to leverage incidental teaching opportunities. Both teachers verbally acknowledged the value of capitalizing on objects or events that capture children's attention (Interview). However, observations showed a significant gap in their implementation capability.

Hale consistently demonstrated the pedagogical flexibility necessary to implement unplanned science activities in response to emergent interest. Her ability to teach opportunistically stemmed from her readily available content knowledge, allowing her to connect phenomena to core concepts in real-time.

Hale noticed that the leaves of one of two identical plants were wilting. She immediately gathered the children around and initiated an unplanned investigation, asking them a series of open-ended, diagnostic questions. A child pointed to the dry soil and stated, "That one looks sad because it's thirsty!" (Observation). Hale used this spontaneous contribution to structure a comparative experiment on plant needs, allowing the children to discover the wilting

cause themselves, thus maintaining child agency in the scientific process.

Conversely, throughout the 13-week observation period, Leyla failed to capitalize on any emergent learning moments, rigidly adhering to her scheduled script. Her limited CK seemed to result in cognitive avoidance, preventing her from engaging with spontaneous inquiries that might lead to unknown conceptual territory.

Leyla picked up a fallen aloe vera leaf and showed the liquid coming out to the children. The plant instantly attracted the children's attention; several children rushed forward, asking, "Is it medicine? Is it juice?" (Observation). Instead of seizing this moment to discuss plant function, moisture, or adaptation, Leyla simply left the leaf with them and walked away. This inability to transition from a behavioral description (showing the liquid) to a conceptual inquiry highlights how low CK restricts pedagogical improvisation.

This marked difference emphasizes that content knowledge in this context is not merely about having information, but about possessing the cognitive capacity to improvise instruction and manage the inherent uncertainty of an inquiry-based environment

Topics

Hale's topic selection was driven by a desire to "stir the curiosity of the children," prioritizing subjects that would support the development of scientific process skills and concepts adaptable to their level (Interview). Leyla, conversely, prioritized familiar concepts and events children encounter in their daily lives, aligning with a lower-risk instructional approach (Interview).

Both teachers covered physical science, Earth and space science, and life science. However, a pattern of competence-based specialization and delegation emerged: Leyla primarily covered physical science (often as demonstrations), while life science topics were consistently delegated to parents. This suggests a strategic avoidance of life science, where her self-reported inadequacy was highest. Hale allocated proportionally more coverage to Earth and space science, reflecting her broad comfort level.

The most significant distinction lay in the teachers' self-efficacy regarding potential conceptual challenges. Hale reported feeling proficient in science subjects and stated she was not worried about children asking unanswerable questions (Interview). This high self-efficacy is a cognitive factor that allowed her to teach more opportunistically.

Leyla, however, admitted to feeling inadequate concerning plants, animals, and physical science, stating directly that she worries when faced with

potential conceptual questions (Interview). While both teachers claimed they would investigate unknown questions, observations indicated Leyla's anxiety led to clear instances of avoiding critical conceptual engagement. Leyla's anxiety and low self-efficacy translated into a risk management strategy of ignoring complex or spontaneous inquiry, thereby protecting her perceived competence level.

During a 'sinking/non-sinking tangerine experiment' (conducted as a demonstration), a child asked a conceptually demanding question: "The ships do not sink because they are big. Is that so, teacher?" Leyla bypassed this question entirely and proceeded with the activity without answering or acknowledging the child's complex hypothesis (Observation). This behavioral pattern demonstrates that Leyla's lack of knowledge inhibited her ability to provide satisfactory answers, directly limiting children's questioning skills and discouraging the natural development of scientific inquiry.

Hale systematically integrated science activities with literacy and music, frequently using picture books to deepen conceptual understanding (Observation). Leyla mostly presented science as a fragmented and isolated activity (Observation). Despite the national curriculum's emphasis on using the environment, both teachers conducted all science activities indoors and did not utilize the outdoor space (Observation). This highlights a persistent systemic/contextual limitation within the school setting that transcends individual teacher competence.

Scientific process skills

The two teachers exhibited starkly different approaches to fostering scientific process skills. Hale actively nurtured skills such as prediction, recall, association, cause-and-effect reasoning, and inference, though she supported classification and measurement less frequently. Leyla's practices, constrained by her self-reported lack of knowledge, generally failed to create an environment conducive to skills development. Notably, neither teacher supported the crucial step of data recording after experiments.

Hale utilized questioning and material manipulation to move children from mere observation to structured inquiry. Her practice was characterized by intentional sequencing to develop measurement skills and foster conceptual discovery.

Hale held a white and a brown egg, initiating a conversation on their similarities and differences. After placing the eggs in the children's hands to facilitate tactile observation, she asked, "They feel different, but how can we know for sure which one is heavier?". A child responded, "We can weigh them, like the flour we used for cookies!". The class then transitioned to using the scale in the science center, demonstrating how Hale's questions successfully prompted children

to propose a testable measurement method based on prior experience (Observation).

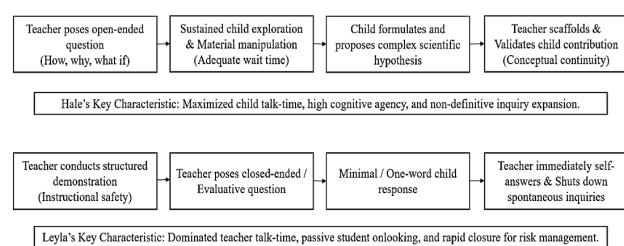
Leyla's instructional style was characterized by a tendency toward conceptual closure and immediate delivery of answers, often bypassing the necessity of active child exploration. This behavior is directly linked to her low self-efficacy, serving as a defense mechanism against spontaneous inquiry. In activities designed to teach scientific concepts, Leyla often employed a "telling" methodology instead of a "discovery" one. In an activity teaching 'hot, warm, and cold' concept, Leyla stated the outcome before the process: "Look, children, this water is a little hot, this water is cold, this water is warm. Now you will come one by one and touch the waters." (Observation). This approach prevents children from formulating hypotheses or engaging in self-regulation of the experiment, reducing the activity to passive sensory exercise.

Furthermore, Leyla demonstrated a lack of pedagogical content knowledge that resulted in the transmission of misinformation. Leyla's attempt to use an analogy to teach freezing failed due to a fundamental conceptual error, which she reinforced by avoiding an experiential follow-up. During a discussion on water melting/freezing, Leyla claimed: "When I put the water in the freezer, it froze. Does everything I put in the freezer freeze? For example, cheese, olives, and jam do not freeze." (Observation). Not only did Leyla fail to wait for children's answers, but she also did not plan to put these materials in the freezer to allow for observation. Most critically, she provided incorrect information by asserting that the water contained in these foods would not freeze, reinforcing a conceptual misunderstanding among the children due to her own lack of sufficient content knowledge.

In comparison, while both teachers included scientific process skills in their plans, Hale consistently created environments that scaffolded true scientific inquiry and discovery. On the other hand, Leyla's need for certainty led her to prioritize instructional efficiency over the development of inquiry skills.

Figure 2.

Behavioral flow models of classroom interaction sequences



Leyla predominantly employed the demonstration experiment technique, which she perceived as the default method for science: "in fact, we always do experiments according to the subject, everything is an experiment" (Interview). This overuse suggests an over-reliance on a structured format that minimizes spontaneous inquiry. Hale, possessing broader content competence, utilized a wider range of methods, integrating more discussion and exploratory activities, even for topics like animals and plants. Furthermore, Hale consciously built conceptual continuity by guiding children to connect current topics to previous ones, a practice Leyla completely omitted.

A stark difference emerged in the level of child engagement permitted during experiments. Hale frequently created an environment that supported active experimentation and true inquiry. In contrast, Leyla relied on demonstration experiments in which children were positioned merely as onlookers.

Leyla's preference for demonstrations extended to experiments that were not only safe but could have been easily adapted for active, hands-on learning (e.g., the non-sinking orange/tangerine experiments). This choice is interpreted as a pedagogical control strategy; by maintaining control over the materials, Leyla limited the potential for unpredictable outcomes or complex spontaneous questions that might expose her conceptual gaps. She even conducted potentially unsafe experiments (e.g., those involving candles, fast reactions) without appropriate safety precautions or conceptual grounding.

The most compelling difference lay in the teachers' use of questioning. Hale consistently employed open-ended questions (e.g., 'how,' 'why,' and 'what if') and practiced excellent scaffolding skills. She ensured children received adequate wait time to formulate complex responses, provided tips, and ensured active participation.

Before reading a book on rain, Hale asked children about the rainy weather. After establishing the fact that it was raining, she moved the inquiry to reasoning by asking, "How did you understand this and what did you see outside?". Later, to explain water vapor, she instructed them to breathe onto the windowpane, using the resulting condensation as a tangible, shared concept (Observation). Hale's practice exemplifies how high pedagogical content knowledge enables the teacher to manage the cognitive load of the lesson while promoting child agency.

In sharp contrast, Leyla relied predominantly on closed-ended questions, offered insufficient time for children to think, and often answered her own questions. This pattern reflects Leyla's lower CK, which increases her own cognitive load during instruction, leading her to avoid demanding, open-ended discourse.

Leyla attempted to teach measurement by placing her foot next to a child's foot and asking the class: "Which one of us has a longer foot? Mine is longer, isn't it?". A child responded, "Yours is longer because you are older?". Leyla immediately affirmed her own statement ("Mine is longer, isn't it?"), shutting down child's attempt to introduce the complex variable of age (Observation). Leyla's closed questioning and quick self-answering prioritized compliance and speed over the development of critical thinking and hypothesis generation.

In summary, Hale utilized a combination of teacher-centered and child-centered methods to create active, inquiry-based environments. Leyla, constrained by her low self-efficacy and limited knowledge, relied almost exclusively on teacher-centered methods and demonstration, prioritizing instructional safety over conceptual depth.

Difficulties encountered

This sub-theme details the challenges experienced by the teachers, which can be categorized into shared pedagogical challenges (classroom environment/transitions) and competence-specific challenges (content adaptation/integration).

Both teachers exhibited difficulties in attracting and maintaining children's attention and ensuring smooth transitions between activities. These challenges were largely linked to structural issues: materials were often prepared only after children had settled for the activity, leading to immediate disengagement. Furthermore, the lack of transition activities, such as finger plays, songs, or rhymes, resulted in recurring classroom management problems for both Hale and Leyla.

The teachers' individual difficulties were directly linked to their core competency levels. Hale's difficulty sometimes arose from her attempt to provide too much detail or over-adapt complex scientific information to the children's level. This led to cognitive overload for children. Hale's high content knowledge sometimes resulted in pedagogical overreach, where her detailed, accurate explanation surpassed the young children's capacity for sustained abstract reasoning.

Explaining the 'bouncing ping-pong ball experiment,' Hale provided a lengthy, multi-step explanation: "You said there was air between this ball and the bottle... When I quickly dropped the ball into the water, there was air in the space under the ball, and thanks to the water flowing into the container, the air there bounced the ball up." Hale repeated this detailed explanation several times, but the children began to get distracted and lose focus (Observation).

Leyla's difficulties were structural, involving a general failure to create an environment that fostered scientific process skills and, critically, an inability to integrate

science with other domains. Her activities were observed as being randomly sequenced and lacking any pedagogical bridge. Leyla's activity sequencing demonstrated a significant lack of pedagogical coherence. The juxtaposition of activities (motor game, experiment, math worksheet) without any intentional connection suggests instruction driven by a checklist rather than a unifying conceptual goal or theme.

"First, Leyla had children play a motor game... Afterward, a 'balloon popping experiment with orange peel' was carried out. Finally, the children painted worksheets with five monkeys and the number 5 (Observation)."

Moreover, Leyla's most profound difficulty was that her limitations in scientific knowledge frequently led to incorrect or incomplete explanations during the experiment, directly compromising the accuracy of the content delivered. This problem was pervasive across activities where she lacked prior conceptual preparation.

Children's reactions during science activities

The final findings provide compelling evidence of how the teachers' pedagogical competence directly shapes the quality of children's participation—a core research question of this study.

Children in Hale's class exhibited active participation, sustained attention, and frequent questioning. This active role was directly influenced by Hale's use of open-ended questions, her provision of sufficient wait time and hints, and her consistent validation of children's responses. This creates a classroom culture of intellectual safety, encouraging children to take risks in inquiry. Hale actively cultivated conceptual continuity by asking questions that required the children to retrieve and synthesize prior knowledge.

In a discussion about eggs, Hale built upon the children's initial answers ("Birds and butterflies, we learned the other day") to pose a more complex, inferential question: "There was a living being in ancient times, but it has not survived to this day. What was its name?" The children successfully retrieved the concept ("dinosaur"), with one child justifying their answer: "Teacher, they had eggs, too. You read us a book; we saw it there." (Observation). This dialogue demonstrates how Hale's scaffolding facilitates knowledge synthesis and metacognition (the ability to reflect on how they know something).

However, consistent with previous observations, Hale's occasional difficulty in adapting complex information led to temporary disengagement (as exemplified by the detailed 'bouncing ping-pong ball experiment' explanation), causing children to become distracted (Observation).

Children in Leyla's class were notably more passive, responding less frequently, asking fewer questions, and exhibiting shorter attention spans. This diminished engagement is attributed to Leyla's limiting behaviors, such as answering her own questions and failing to follow up on children's inquiries. Crucially, Leyla's limitations in scientific knowledge not only hampered participation but also directly contributed to the development and reinforcement of scientific misconceptions among the children.

During a hot/cold/warm water demonstration, Leyla focused on safety by telling children to stay away from the table due to the hot water. When a child observed the rising steam and hypothesized, "It must be really hot; there is smoke coming out," Leyla failed to intervene (Observation). This inaction represents a missed opportunity to correct the common misconception that steam is smoke—a failure stemming from either a lack of conceptual vigilance (PCK deficit) or a prioritization of classroom management over conceptual clarity.

This finding strongly suggests that Leyla's instructional deficiencies created a low-risk, low-gain learning environment, discouraging participation and inadvertently allowing scientifically inaccurate beliefs to be validated through teacher silence.

Discussion and Conclusion

The purpose of this research was to compare the planning and implementation of science activities by two kindergarten teachers: one with a strong background in scientific knowledge and a positive attitude toward science education, and the other with considerably less expertise. The results indicate that the teachers had differences in both the planning and implementation processes.

The teacher with stronger scientific knowledge demonstrated a systematic approach to planning, actively researching topics and gathering children's literature and resources. In contrast, the teacher with limited science knowledge relied primarily on demonstration experiments sourced from the internet, often without a deeper understanding of the concepts involved. This difference can be interpreted through the lens of pedagogical content knowledge and teacher self-efficacy theories (Bandura, 1997; Gudmundsdottir & Shulman, 1987). Teachers with insufficient content knowledge may experience reduced self-efficacy, leading them to adopt safer, more structured teaching methods—such as demonstrations—where the potential for conceptual error is minimized. Similar observations have been made in previous research, indicating that teachers' limited understanding of science inhibits their willingness to explore inquiry-based learning environments (Bilaloğlu et al., 2008; Greenfield et al., 2009; Tu, 2006).

A particularly original finding of this study concerns how the teacher with limited expertise compensated for her lack of scientific knowledge by involving parents. She delegated science activities to the "parent of the week," who conducted them during class. This practice, while increasing the number of activities, did not ensure pedagogical coherence or inquiry-oriented interaction. However, it offers valuable insight into how early childhood teachers in Türkiye may navigate professional challenges. From a sociocultural perspective, this behavior reflects adaptive compensation—teachers drawing on family participation as a contextual resource to overcome their own professional limitations. In many Western contexts, family engagement is conceptualized as a collaborative partnership (Epstein & Miller, 2011), whereas in this study it functioned as a compensatory mechanism for pedagogical gaps. Consequently, the 'parent-teacher-child' triangle in the Turkish educational ecosystem structurally diverges from Western frameworks by repositioning the parent from an external learning partner into an active instructional substitute. This specific cultural adaptation highlights how intense parental expectations and institutional pressures in a centralized system can prompt teachers to utilize family resources to manage their own professional anxieties and content knowledge deficits. This finding fills a notable gap in the international literature by illustrating how cultural expectations around family participation can intersect with teacher competencies and instructional practices.

The relationship between teachers' science knowledge and their questioning techniques also emerged as a critical finding. The teacher with stronger science knowledge frequently used open-ended questions (why, how, and what if) that stimulated reasoning, while the teacher with limited knowledge relied predominantly on closed-ended questions. This pattern can be explained by cognitive load theory: when teachers lack sufficient conceptual understanding, their working memory is occupied with managing the activity itself, leaving limited capacity for spontaneous inquiry (Sweller, 2011). As a result, they avoid cognitively demanding questioning. This finding reinforces earlier reports that 95% of early childhood teachers' questions tend to be closed-ended (Bilaloğlu et al., 2017; Siraj-Blatchford & Manni, 2008), and that teachers' questioning quality is strongly linked to their confidence in subject matter knowledge.

Furthermore, the teacher with stronger science knowledge demonstrated an ability to respond to unplanned "teachable moments" by recognizing children's spontaneous questions and extending them into meaningful inquiry. Such responsiveness is central to reflective practice (Schön, 2017) and inquiry-based learning, both of which require teachers to continuously adapt instruction based on children's

curiosity and prior experiences (McCray & Chen, 2012; Oppermann et al., 2021). The other teacher's inability to capitalize on these moments highlights the pedagogical limitations that arise when content knowledge is shallow and rigidly scripted.

Integration across domains also distinguished the two teachers. The teacher with higher science competence connected science activities to literacy, art, and mathematics, which aligns with research showing that interdisciplinary approaches enhance children's cognitive flexibility and holistic development (Bevan et al., 2020; Pella, 2022). In contrast, the other teacher's activities remained fragmented and isolated. This suggests that understanding the interconnectedness of disciplines is not merely desirable—it is foundational to effective early childhood science education (Edwards & Loveridge, 2011; French, 2004).

Another critical finding was that the teacher with limited science knowledge provided inaccurate explanations, particularly in physical sciences. Misinterpretations of scientific phenomena—such as using demonstrations without explaining underlying principles—can foster misconceptions that persist into later education (Binns & Popp, 2013; Gerde et al., 2018; Kallery & Psillos, 2001). Children's misconceptions often arise not from curiosity but from teachers' oversimplifications or conceptual errors. Hence, teacher preparation programs should focus not only on what to teach but also on how to represent concepts in developmentally appropriate ways (Akyol & Birinci-Konur, 2018; Dağlı & Dağlıoğlu, 2020).

Several methodological limitations should be acknowledged. First, the study relied on observations of two teachers, and video recording was not permitted. Second, the researcher was the sole observer, which may have introduced interpretive bias. Third, the observation period was limited. Fourth, because this investigation was situated exclusively within private preschools in a major metropolitan center (Istanbul), the findings predominantly illustrate urban early childhood dynamics. The enacted pedagogical content knowledge, resource availability, and classroom science practices may manifest differently among educators working in rural, remote, or socio-economically disadvantaged regions of Türkiye, thereby bounding the immediate transferability of these outcomes to non-metropolitan environments. Nevertheless, the study remains valuable because it provides rich qualitative insights into authentic classroom practices that are rarely captured in early childhood science education research in Türkiye. Despite these constraints, the findings offer a nuanced understanding of how teachers with differing science backgrounds enact, adapt, and compensate within real educational contexts. To strengthen early childhood science education, several actionable steps can be derived from these findings:

Undergraduate early childhood programs in Türkiye currently offer only one science education course. This structural brevity creates a substantial gap when contrasted with international best practices and professional standards. For instance, the National Association for the Education of Young Children (NAEYC) guidelines emphasize that high-quality teacher preparation must integrate continuous, multi-semester science and STEM pedagogy interwoven with active field practicums (NAEYC, 2020). Integrating foundational physical and biological science courses—similar to those in the Missouri State University Early Childhood Program (2025)—can build stronger conceptual understanding and practical confidence. Restructuring the national curriculum to shift from a single-course model to a consecutive, inquiry-based science methodology track is imperative to equip pre-service teachers with sustainable Pedagogical Content Knowledge (PCK).

In-service training should move beyond content delivery to include mentoring, microteaching, and reflective feedback components that develop teachers' inquiry-based instructional skills.

Given the observed reliance on family support, schools should design structured, pedagogically guided family engagement models that treat parents as learning partners rather than substitutes for teacher expertise.

Although the Turkish Early Childhood Education Program (MoNE, 2024) emphasizes outdoor learning, both teachers confined activities indoors. Future training should include modules on using outdoor spaces for inquiry-based exploration (Loxley et al., 2016).

In conclusion, this study demonstrates that teachers' levels of scientific knowledge and attitudes toward science directly influence the frequency, quality, and pedagogical richness of science activities in early childhood education. Teachers with stronger scientific backgrounds fostered more inquiry-oriented, integrated, and responsive learning environments, whereas those with limited knowledge depended on structured demonstrations and external support. Importantly, the finding that teachers compensate for their lack of knowledge through family involvement contributes a novel cultural and pedagogical perspective to the global literature on teacher competence. Addressing these challenges through enriched coursework, reflective professional learning, and structured family partnerships will be crucial to advancing the quality and equity of early childhood science education in Türkiye and beyond.

AI Disclosure Statement

During the preparation of the manuscript, the authors used Google Gemini to assist with English translation

and language editing, including improvements in grammar, wording, and readability. All AI-assisted translations and revisions were carefully reviewed, verified, and edited by the authors, who take full responsibility for the accuracy and content of the final manuscript.

References

- Aktaş, M. (2024). Digital transformation in early childhood education: Evaluating hybrid learning tools in preschool science activities. *Journal of Early Childhood Research*, 22(2), 145–162. <https://doi.org/10.1177/1476718X241234567>
- Akyol, N., & Birinci-Konur, K. (2018). The examination of preschool teachers' and school managers' views on the applicability of pre-school science education. *Kastamonu Education Journal*, 26(2), 547–557. <https://dergipark.org.tr/pub/kefdergi/issue/31580>
- Babaroğlu, A., & Okur-Metwalley, E. (2018). Opinions of preschool teachers on science education in early childhood. *Hitit University Journal of Social Sciences Institute*, 11(1), 125–148. <https://doi.org/10.17218/hititsosbil.389149>
- Bandura, A. (1997). *Self-efficacy: The exercise of control*. Freeman.
- Barenthien, J., Lindner, M. A., Ziegler, T., & Steffensky, M. (2020). Exploring preschool teachers' science-specific knowledge. *Early Years*, 40(3), 335–350. <https://doi.org/10.1080/09575146.2018.1443321>
- Bevan, B., Gutwill, J. P., Petrich, M., & Wilkinson, K. (2020). Learning through STEM-rich tinkering: Findings from a jointly negotiated research project taken to scale. *Science Education*, 104(3), 500–525. <https://doi.org/10.1002/sce.21576>
- Bilaloğlu, R. G., Aktaş-Arnas, Y., & Yaşar, M. (2017). Question types and wait-time during science related activities in Turkish preschools. *Teachers and Teaching*, 23(2), 211–226. <https://doi.org/10.1080/13540602.2016.1203773>
- Bilaloğlu, R. G., Aslan, D., & Aktaş-Arnas, Y. (2008). The analysing of preschool teachers' levels of knowledge about science activities. *The Journal of National Education*, 37(178), 88–104. <https://dergipark.org.tr/pub/milliegitim>
- Binns, I. C., & Popp, S. (2013). Learning to teach science through inquiry: Experiences of preservice teachers. *The Electronic Journal for Research in Science & Mathematics Education*, 17(1), 1–24.
- Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77–101. <https://doi.org/10.1191/1478088706qp0630a>
- Cho, H., Kim, J., & Choi, D. H. (2003). Early childhood teachers' attitudes toward science teaching: A scale development study. *Journal of Early Childhood Education*, 23(4), 213–231.
- Creswell, J. W., & Creswell, J. D. (2018). *Research design* (5th ed.). SAGE.
- Dağlı, H., & Dağlıoğlu, H. E. (2020). An analysis of preschool teachers' opinions on the content and standards of science education. *OPUS International Journal of Society Researches*, 15(23), 1885–1919. <https://doi.org/10.26466/opus.631378>
- Edwards, K., & Loveridge, J. (2011). The inside story: Looking into early childhood teachers' support of children's scientific learning. *Australian Journal of Early Childhood*, 36(2), 28–35. <https://doi.org/10.1177/183693911103600205>
- Epstein, D., & Miller, R. T. (2011). *Slow off the mark: Elementary school teachers and the crisis in science, technology, engineering, and math education*. Center for American Progress. <https://eric.ed.gov/?id=ED536070>
- French, L. (2004). Science as the center of a coherent, integrated early childhood curriculum. *Early Childhood Research Quarterly*, 19(1), 138–149. <https://doi.org/10.1016/j.jecresq.2004.01.004>
- Genç-Kumtepe, E., Kaya, S., & Kumtepe, A. T. (2009). The effects of kindergarten experiences on children's elementary science achievement. *Elementary Education Online*, 8(3), 978–987. <https://www.ilkogretim-online.org/?mno=121728>
- Genç-Kumtepe, E., Kumtepe, A. T., & Batmaz, B. (2013). The convergence of perceived efficacy beliefs and sciencing in early childhood classrooms. *Mustafa Kemal University Journal of Social Sciences Institute*, 10(21), 85–99. <https://dergipark.org.tr/en/pub/mkusbed/issue/19547/208171>
- Gerde, H. K. J., Pierce, S., Lee, K., & Van-Egeren, L. A. (2018). Early childhood educators' self-efficacy in science, math, and literacy instruction and science practice in the classroom. *Early Education and Development*, 29(1), 70–90. <https://doi.org/10.1080/10409289.2017.1360127>
- Gray, D. E. (2004). *Doing research in the real world*. SAGE Publications.

- Greenfield, D. B., Jirout, J., Dominguez, X., Greenberg, A., Maier, M., & Fuccillo, J. (2009). Science in the preschool classroom: A programmatic research agenda to improve science readiness. *Early Education and Development*, 20(2), 238–264. <https://doi.org/10.1080/10409280802595441>
- Gudmundsdottir, S., & Shulman, L. (1987). Pedagogical Content Knowledge in Social Studies. *Scandinavian Journal of Educational Research*, 31(2), 59–70. <https://doi.org/10.1080/0031383870310201>
- Güney, M. Y. (2016). *The effect of activities based upon constructivism on preschool teacher candidates' attitudes and achievements* [Unpublished master's thesis]. Kastamonu University.
- Kallery, M., & Psillos, D. (2001). Pre-school teachers' content knowledge in science: Their understanding of elementary science concepts and of issues raised by children's questions. *International Journal of Early Years Education*, 9(3), 165–179. <https://doi.org/10.1080/09669760120086929>
- Loxley, P., Dawes, L., Nicholls, L., & Dore, B. (2016). *Teaching primary science: Promoting enjoyment and developing understanding* (3rd ed.). Routledge.
- McCray, J. S., & Chen, J. Q. (2012). Pedagogical content knowledge for preschool mathematics: Construct validity of a new teacher interview. *Journal of Research in Childhood Education*, 26(3), 291–307. <https://doi.org/10.1080/02568543.2012.685123>
- Mercan, N., & Kutluca, A. Y. (2021). Investigation of the effect of preschool teachers' epistemological beliefs on their pedagogical competencies for science teaching. *Mehmet Akif Ersoy University Journal of Education Faculty*, 59, 30–59. <https://dergipark.org.tr/en/pub/maeuefd/issue/64363/852671>
- Merriam, S. B. (2018). *Qualitative research: A guide to design and implementation* (3rd ed.). Jossey-Bass.
- Miles, M. B., & Huberman, A. M. (1994). *Qualitative data analysis: An expanded sourcebook*. SAGE.
- Ministry of National Education. (2016). *TIMSS 2015 national math and science preliminary report*. http://timss.meb.gov.tr/meb_iys_dosyalar/2022_03/07135609_TIMSS_2015_Ulusal_Rapor.pdf
- Ministry of National Education. (2024). *Preschool education program: Türkiye's Century Education Model*. <https://tymm.meb.gov.tr/ogretim-programlari/okul-oncesi>
- Missouri State University. (2025). *BSEd four-year degree plan: Early childhood education*. <https://education.missouristate.edu/TLDS/ECE/FourYearPlan.htm>
- Morgan, P. L., Farkas, G., Hillemeier, M. M., & Maczuga, S. (2016). Science achievement gaps begin very early, persist, and are largely explained by modifiable factors. *Educational Researcher*, 45(1), 18–35. <https://doi.org/10.3102/0013189X16633182>
- National Association for the Education of Young Children. (2020). *Professional standards and competencies for early childhood educators*. <https://www.naeyc.org/resources/position-statements/standards-competencies>
- Nilsson, P., & Loughran, J. (2012). Exploring the development of pre-service science teachers' pedagogical content knowledge: The role of prompts in formative assessment. *International Journal of Science Education*, 34(13), 1965–1990. <https://doi.org/10.1080/09500693.2012.684114>
- O'Connor, G., Fragkiadaki, G., Fleer, M., & Rai, P. (2021). Early childhood science education from 0 to 6: A literature review. *Education Sciences*, 11(4), 178. <https://doi.org/10.3390/educsci11040178>
- Olğan, R. (2015). Influences on Turkish early childhood teachers' science teaching practices and the science content covered in the early years. *Early Child Development and Care*, 185(6), 926–942. <https://doi.org/10.1080/03004430.2014.967689>
- Oppermann, E., Hummel, T., & Anders, Y. (2021). Preschool teachers' science practices: Associations with teachers' qualifications and their self-efficacy beliefs in science. *Early Child Development and Care*, 191(5), 800–814. <https://doi.org/10.1080/03004430.2019.1647191>
- Organisation for Economic Co-operation and Development. (2019, March 8). *PISA 2018 results: Combined executive summaries volume I, II & III*. https://www.oecd.org/pisa/Combined_Executive_Summaries_PISA_2018.pdf

- Organisation for Economic Co-operation and Development. (2022). *PISA 2022 results: Comparing countries' and economies' performance*. https://www.oecd.org/pisa/OECD_2022_PISA_Results_Comparing%20countries%E2%80%99%20and%20economies%E2%80%99%20performance%20in%20mathematics.pdf
- Öztürk, F. (2025). Post-pandemic pedagogical adaptations: Techno-pedagogical content knowledge among Turkish kindergarten teachers. *Eurasian Journal of Educational Research*, 112, 89–110. <https://doi.org/10.14689/ejer.2025.112.005>
- Park, S., & Oliver, J. S. (2008). Revisiting the conceptualisation of pedagogical content knowledge (PCK): PCK as a conceptual tool to understand teachers as professionals. *Research in Science Education*, 38(3), 261–284. <https://doi.org/10.1007/s11165-007-9049-6>
- Pella, S. (2022). Interdisciplinary teaching and learning in early childhood education: A systematic review. *Early Childhood Education Journal*, 50(1), 1–14. <https://doi.org/10.1007/s10643-020-01112-7>
- Pendergast, E., Lieberman-Betz, R. G., & Vail, C. O. (2017). Attitudes and beliefs of prekindergarten teachers toward teaching science to young children. *Early Childhood Education Journal*, 45, 43–52. <https://doi.org/10.1007/s10643-015-0761-y>
- Polat, S., Zengin, R., & Elmalı, F. (2021). Attitude and practice analysis of preschool teachers for science practice in classroom. *Turkish Journal of Educational Studies*, 8(1), 47–67. <https://dergipark.org.tr/en/pub/turkjes/issue/60520/771896>
- Roberts, P., & Collins, P. R. (2026). Praxis: The Missing Theory to Practice Link in Maximising Early Childhood Science Opportunities. *Education Sciences*, 16(1), 134. <https://doi.org/10.3390/educsci16010134>
- Sağkes, M. (2014). How often do early childhood teachers teach science concepts? Determinants of the frequency of science teaching in kindergarten. *European Early Childhood Education Research Journal*, 22(2), 169–184. <https://doi.org/10.1080/1350293X.2012.704305>
- Schön, D. A. (2017). *The reflective practitioner: How professionals think in action*. Routledge. <https://doi.org/10.4324/9781315237473>
- Simsar, A., & Davidson, S. G. (2020). Sources of preservice early childhood teachers' self-efficacy beliefs about teaching science: A phenomenological study. *Eğitimde Nitel Araştırmalar Dergisi*, 8(3), 776–795. <https://izlik.org/JA73DE47MG>
- Snape, D., & Fox, J. (2021). Parental expectations and institutional pressures in early years education: A cross-cultural analysis. *International Journal of Early Years Education*, 29(4), 382–398. <https://doi.org/10.1080/09669760.2021.1892110>
- Sweller, J. (2011). Cognitive load theory. *Psychology of Learning and Motivation* 55, 37–76. <https://doi.org/10.1016/B978-0-12-387691-1.00002-8>
- Silverman, D. (2015). *Interpreting qualitative data* (5th ed.). SAGE Publications.
- Siraj-Blatchford, I., & Manni, L. (2008). Would you like to tidy up now? An analysis of adult questioning in the English Foundation Stage. *Early Years*, 28(1), 5–22. <https://doi.org/10.1080/09575140701842213>
- Tu, T. (2006). Preschool science environment: What is available in a preschool classroom?. *Early Childhood Education Journal*, 33(4), 245–251. <https://doi.org/10.1007/s10643-005-0049-8>
- Ünal, M. P., Akman, B., & Gelbal, S. (2010). The adaptation of a scale for preschool teachers' attitudes towards science teaching. *Procedia - Social and Behavioral Sciences*, 2, 2881–2884. <https://doi.org/10.1016/j.sbspro.2010.03.433>
- Yin, R. K. (2018). *Case study research and applications: Design and methods* (6th ed.). SAGE Publications.



This page is intentionally left blank.
www.iejee.com

Using Social Media for Participant Recruitment in Early Childhood Education and Care Research: A Reflective Analysis

Elysia Phillips^{a*}, Margaret Sims^b, Nicola Schutte^c, Marg Rogers^d

Received : 22 December 2025
Revised : 28 June 2026
Accepted : 18 July 2026
DOI : 10.26822/iejee.2026.464

^a**Corresponding Author:** Elysia Phillips
School of Education, University of New England,
Armidale, Australia
E-mail: ephilips@une.edu.au
ORCID: <https://orcid.org/0009-0005-8166-4515>

^b Margaret Sims, Macquarie University, Australia.
E-mail: margaret.sims@mq.edu.au
ORCID: <https://orcid.org/0000-0003-4686-4245>

^c Nicola Schutte, School of Psychology, University of
New England, Armidale, Australia.
E-mail: nschutte@une.edu.au
ORCID: <https://orcid.org/0000-0002-3294-7659>

^d Marg Rogers, School of Education, University of
New England, Armidale, Australia ; Manna Institute,
Australia.
E-mail: marg.rogers@une.edu.au
ORCID: <https://orcid.org/0000-0001-8407-7256>

Abstract

This methodological reflection examines the practical and ethical experiences of applying a constructionist methodological approach in a higher degree research project exploring Australian early childhood educators' understanding of emotional intelligence (EI) and how it relates to their daily practice. Three key reflections emerged: navigating ethics approval, using social media as a recruitment tool and managing micro-ethical dilemmas related to geo-location and participant interactions in online environments. These challenges highlight broader issues identified in literature regarding ethical complexity, researcher reflexivity and navigation of digital spaces (Buchanan & Zimmer, 2012). The online ethics application presented additional layers of procedural nuance, providing an opportunity for professional learning. The use of social media was advantageous in reaching a wide range of early childhood education and care workforce, aligning with literature that digital recruitment can enhance accessibility (Gelinas et al., 2017) however complicated the intended scope and raised issues of clarity of selection criteria and management of participants. Given the content of the survey addressed sensitive topics around emotional strain, stress and burnout, the research underscored the importance of maintaining transparency and reflexivity when interacting in digital spaces. These insights contribute to existing discussions about ethical practice in early childhood research and the digital research context, particularly for novice researchers and research supervisors.

Keywords:

Methodological Reflection, Ethics, Participant Recruitment, Partnership for the Goals, Social Media, Data Collection, Quality Education, Early Childhood Educators, Professionalism



www.iejee.com
ISSN: 1307-9298

2026 Published by KURA Education & Publishing.
This is an open access article under the CC BY-NC-ND license. (<https://creativecommons.org/licenses/by/4.0/>)

Introduction

Methodological choices play a central role in shaping research design, data collection and interpretation of results, particularly when exploring complex social and emotional challenges (Doyle, 2009; Lyndon, 2023). Beyond the technical considerations, research methodology is increasingly understood as a reflexive process in which researchers critically examine how their decisions,

assumptions and interactions influence the research process and the outcomes of the study (Berger, 2015; Finlay, 2002). Therefore, methodological reflections serve as an important mechanism for the evaluation of research practices, supports in identifying challenges and generating insights that have the potential to inform future research.

As research has become increasingly performed within digitally mediated environments, the importance of methodological reflections has become more evident. Digital technologies and social media platforms have expanded opportunities for participant recruitment, communication, and dissemination of findings, offering researchers access to more geographical and broader populations. These technologies can increase recruitment efficiency and reduce costs to conduct research studies while facilitating engagement with professional communities that may be difficult to access through more traditional measures. However, with this increase in the use of digital spaces in research, it introduces new methodological and ethical considerations relating to privacy, confidentiality, recruitment, informed consent and clear distinctions between public and private (Franzke et al., 2020; Markham & Buchanan, 2012).

Ethical considerations within research extend beyond formal institutional approval processes. Guillemin and Gillam (2004) argue that while procedural ethics provides an important framework for protecting participants, researchers must all navigate ethics in practice, often the unexpected ethical issues that emerge when conducting research. These micro-ethical dilemmas (Burke et al., 2024) require researchers to exercise reflexive judgement when responding to situations that cannot always be anticipated during the ethics approval process. Such dilemmas may arise at any stage of the research, such as participant recruitment, data collection and the dissemination of findings. The dynamic and changing nature of digital research can further intensify the ethical complexities, requiring ongoing consideration of safeguards to protect both researcher and participants, including issues related to privacy, confidentiality, online interactions and wellbeing.

This paper presents reflections from a study examining Australian early childhood educators (ECEs) experiences of burnout, joy and feelings of being overwhelmed in relation to emotional intelligence (EI). The study utilised online survey methods and social media recruitment strategies to engage Australian educators. Three key reflections from this study focus on navigating ethics approval, using social media for participant recruitment and addressing micro-ethical dilemmas associated with geo-location and the dissemination of findings. These reflections are grounded in practical and ethical challenges

encountered throughout the research process and consider the perspective of the lead author (a higher degree research student) and the supervisory team (co-authors). Through these reflections, the authors explore the methodological decisions, ethical complexity, and institutionalised system that influence research processes and outcomes. These reflections offer insights that may inform future research conducted within educational and digitally mediated contexts.

Early Childhood Education and Care (ECEC) Context

These methodological reflections are situated within the Australian early childhood education and care (ECEC) sector, a workforce characterised by emotional labour, complex interpersonal relationships and increasingly demanding professional expectations (Andrew, 2015; Bull et al., 2024). ECEs are required to navigate relationships with children, families, colleagues, and regulatory authorities while maintaining responsive and supportive learning environments. Recent research has highlighted growing concerns regarding staff retention, educator wellbeing, stress, and burnout across the sector which reinforces the importance of research that seeks to understand educators lived experience (Bull et al., 2024; McMullen et al., 2020). Given these challenges, conducting ethical research in ECEC requires careful consideration of educator wellbeing and the contextual realities of the workforce.

This study was conducted in Australia and involved educators working across a range of ECEC services regulated under the National Quality Framework (NQF). The NQF establishes national standards relating to various components of early childhood, including but not limited to, children's health and hygiene, educational programs and practices, staffing arrangements and leadership (Australian Children's Education and Care Quality Authority (ACECQA), 2019). Educators are geographically dispersed across metropolitan, regional, rural and remote communities creating practical challenges for participant recruitment. Researchers have increasingly recognised that traditional methods such as email distribution lists and professional networks may not always provide sufficient access to the diverse educator population.

Opportunities have been created with the introduction of digital technologies and social media platforms in supporting a diverse range of participant recruitment. Platforms such as Facebook and LinkedIn enable researchers to reach large professional communities quickly and cost effectively, while gaining access to participants across broad geographical locations. Previous studies have reported that social media recruitment can improve recruitment efficiency, increase participant diversity and enhance access to populations that may otherwise be difficult to engage

(Gelinas et al., 2017; Whitaker et al., 2017). There have been limitations identified including self-selection bias, privacy risks, representation concerns and challenges with informed consent within digital spaces (Gelinas et al., 2017; Zimmermann et al., 2022).

These challenges are particularly relevant when recruiting professionals whose online identities often span both personal and professional domains. Social media platforms can blur the boundaries between public and private spaces creating unique ethical complexities regarding participant privacy, researcher viability and informed consent (Zimmermann et al., 2022). Consistent with recommendations for online research, participants in this study were provided with a detailed participant information statement outlining the study's purpose, procedural risks, and participant rights. Consent was embedded within the survey requiring participants to actively indicate agreement before continuing to the questions. This approach was designed to ensure participation remained voluntary and informed, while maintaining anonymity (Early Childhood Australia, 2022).

The significance of these ethical considerations is becoming more recognised worldwide through an increase in the amount of documented ethical guidelines by various state and community bodies. For example, the European Early Childhood Education Research Association (EECERA) emphasises the importance of robust ethical decision making throughout all stages of research, highlighting the need for researchers to respond reflexively to emerging ethical issues within digital spaces (Bertram et al., 2025). Similarly, Early Childhood Australia (2022) advocates for ethical frameworks that acknowledge evolving complexities of digital research within educational settings. Despite the growing use of social media recruitment across disciplines, relatively limited attention has been given to practical and ethical implications within ECEC research.

This paper underscores the need for tailored ethical guidelines that address specific challenges of conducting research within early childhood settings. These reflections are situated within the Australian early childhood education and care sector, providing guidance for novice researchers, supervisors and research institutions seeking to balance methodological severity with ethical responsibility in multifaceted educational contexts.

Ethical Considerations: Dissemination, Approval, Recruitment, and Conduct

Throughout this study, ethical considerations were at the forefront of each stage of development, from ethics approval to disseminating findings. The reflections presented in this article focus on dissemination, however, are part of a broader ethical path that

included approval, recruitment and data collection. This approach aligns with literature highlighting the value of ongoing reflexivity in research practices, especially for higher degree research students (Menning et al., 2025; Zimmermann et al., 2022).

Dissemination of Findings

The research team engaged in systematic reflexivity to ensure the findings in this reflective article were presented ethically and responsibly. Decisions regarding the inclusion of materials were guided by the Association of Internet Researchers (2019) principles. Each potential risk was considered in relation to harm, contextual integrity and confidentiality. For example, a social media comment was retained to illustrate micro-ethical dilemmas; the information is publicly accessible, however, all identifying details were removed.

These decisions were made collaboratively as a team; the lead author proposed the approach, the supervisors reviewed and provided direct guidance and instruction, and the procedures were cross-checked against ethical guidelines. While formal IRB approval was not required, the decision-making process was informed by established ethical principles and reflective dialogue within the research team. This approach ensured that the materials used were ethically and contextually sensitive.

Ethical Rationale

The ethical issues highlighted in dissemination were chosen because they represent the recurring challenges in digital research. Digital platforms can blur the boundaries between public and private spaces, complicating consent and participant autonomy (Zimmermann et al., 2022). To address these concerns, researchers need to embed consent into the survey flow, clearly articulate participant rights and adopt transparent practices (Early Childhood Australia, 2022). Decisions were guided by a principle of situated ethical reflexivity, which weighed the value of including illustrative examples against the responsibility to protect participants and the institution.

By documenting this reflective process, the paper contributes to literature on ethical practice in early childhood education and care research, demonstrating practical strategies for balancing transparency, scholarly rigour and participant anonymity (Bertram et al., 2025; Zimmermann et al., 2022). In addition, these reflections highlight the importance of professional development for novice researchers in navigating ethical complexities.

In the following sections, the lead author reflects on the challenges she encountered and discusses each in relation to the literature. Finally, we draw our

conclusions and provide recommendations to inform policy and practice within research institutions.

Lead Author Reflection: Ethics Approval

The lead author, as a higher degree research (HDR) student, navigated the institutions online ethics approval platform for the first time. The platform was designed to streamline applications. While intended to enhance efficiency, usability challenges were encountered regarding account set-up, understanding the different sections and managing multi-level approval.

These challenges contributed to delays in commencing the study, especially when clarifying the inclusion of an honorarium for early childhood educators who participated in the study, who are typically underpaid and hold low-status roles (Bull et al., 2024). The research team carefully considered the ethical implications of offering compensation, balancing respect for educators' time with avoidance of undue inducement, in alignment with the EECERA Ethical Code (Bertram et al., 2025).

Within the lead author's reflection, the cause of these delays was exacerbated by limited prior training and the need to coordinate with supervisors who were also unfamiliar with the platform. Multiple submissions and requests for clarification were necessary to ensure compliance with ethical standards. This demonstrates that online systems require user familiarity and institutional support for effective implementation.

Team Discussion

While the online platforms are proposed to increase efficiency and transparency within the ethical approval process (Rani et al., 2024; Sheehy et al., 2020), evidence suggests the benefits are not automatically realised. Usability challenges can slow approvals, creating stress for HDR students and requiring supervisory intervention. Dialogue between supervisors and students can mitigate these challenges by facilitating ethical decision-making and micro-reflection on the research process (Burke et al., 2024). Institutions implementing digital platforms need to provide targeted training and ongoing support to ensure novice researchers can effectively navigate the ethical process.

Lead Author Reflection: Social Media Participant Recruitment

Social media recruitment was employed to reach a broader audience of Australian early childhood educators. Initial posts of platforms such as LinkedIn and Facebook inadvertently attracted international participants, despite the study's focus on Australian educators. The post was revised to explicitly state

inclusion criteria and re-shared via Facebook to engage the target population.

Ethical considerations included informed consent, privacy and authenticity of engagement. Embedded in the survey flow was a multi-page information sheet and consent form; only participants who selected "yes" could proceed.

When posting on social media (e.g. LinkedIn and Facebook), the team received some positive feedback from the online community, who emphasised the importance of this study. They commented, "... emotional intelligence is crucial in education. Engaging with this research helps develop supportive environments for our children. Well done promoting such important studies". The contributor's feedback further validated the importance of this research and the relevance of identifying the role EI has in shaping educators' practices and experiences within the ECEC sector (Kapoor et al., 2018). This process of reflection, adjustment, and filtering helped ensure the integrity of the research continued to align with the study's aims and outcomes, thereby achieving its intended scope.

Team Discussion

The use of social media in participant recruitment has become increasingly popular. This is due to its effectiveness and ease of use in recruiting participants for studies and reaching new cohorts (Darko et al., 2022). However, while there are many advantages to using social media, it can also introduce challenges and raise numerous ethical considerations that need to be addressed in the study design and ethics application. Some of the advantages of using social media platforms, such as Facebook and LinkedIn, are that they provide researchers with access to a diverse audience. Researchers can recruit participants from a wide range of diverse demographic backgrounds, making it easier and accessible to find people who match the specific criteria (Darko et al., 2022). Additionally, it is a low-cost method of recruitment, with postings available free of charge; the only requirement is to have an account. Furthermore, social media is a fast way to recruit participants (Gelinas et al., 2017). Social media platforms can enable you to tailor your posts when targeting specific groups based on the study's criteria. It is valuable as it provides a large sample of the population at a low cost (Benedict et al., 2019).

While there are many benefits to the use of social media, it also raises some concerns regarding bias, privacy and confidentiality, among others. Bias may occur as it is not a random sample of the population; however, it is accessible to particular groups, which may affect the generalisation of the research findings (Darko et al., 2022). Recruiting by social media also poses an ethical challenge about privacy, confidentiality and

informed consent (Benedict et al., 2019). For example, the nature of social media may blur the lines between personal interaction and professional research (Gelinás et al., 2017). Additionally, participants may question the authenticity of the research recruitment, given the prevalence of misinformation online. In this study, social media has changed the way researchers recruit participants by providing a powerful tool to engage a broader and more diverse range of people. This approach must address the ethical concerns, such as informed consent, privacy and confidentiality (Gelinás et al., 2017). By following ethical guidelines, researchers can effectively utilise social media to reach a diverse range of participants.

Lead Author Reflection: Ethical Dilemma with Geo-location Data

During the data collection stage via Qualtrics, an online survey platform, geo-location tags revealed that some participants were outside of Australia, despite the eligibility criteria. The lead author manually filtered out responses to include only Australian educators, discarding international responses. This decision was guided by ethical obligations to ensure data integrity and alignment with the approved study scope.

This scenario presented a micro-ethical dilemma (Burke et al., 2024), requiring reflection on balancing ethical principles, participant respect and scholarly rigour. Decisions were documented, as well as supervisors' consultations. As a result, amendments were made to ensure transparency and adherence to ethical codes.

Team Discussion

As HDR students conduct research, they may encounter a range of challenges when navigating research ethics. Challenges and situations arise that require ethical awareness and flexibility in navigating the process without compromising the integrity of the research (Taquette & Borges da Matta Souza, 2022). For example, Burke et al. (2024) describe the term 'micro-ethical dilemma', which refers to the smaller, subtle ethical challenges that may arise during the conduct of the research. These micro-ethical dilemmas extend beyond what was anticipated and recorded in the ethics application process, requiring flexibility from both the HDR student and the supervisory team. Burke et al. (2024) identified dialogue as being the most important way that supervisors and HDR students can work through these challenges. Dialogue creates the opportunity for the student researcher to develop micro-ethical reflexivity, which is a skill combining ethical awareness in the moment and later reflecting on how it impacted their actions and decisions as a researcher (Burke et al., 2024).

Integrating Literature and Reflection

The reflections presented in this study highlight the complexities of conducting research as a HDR student, particularly within the early childhood education and care sector, where ethical considerations are amplified by the emotional labour and relational nature of the professional practice (Andrew, 2015; Bull et al., 2024; Mérida-López & Extremera, 2017). By integrating the practical experience of the lead author within the existing literature, this section demonstrates how micro-ethical dilemmas, methodological decisions and institutional systems intersect to shape research outcomes.

Ethics Approval and Digital Platforms

The lead author's experience with the institution's online ethics platform illustrates the tension between institutional narratives of efficiency and the lived experience of first-time researchers. While digital ethics systems have been designed to streamline applications, increase transparency and reduce administrative burden (Rani et al., 2024; Sheehy et al., 2020), these benefits are not automatically realised. As noted by Van De Mierop et al. (2022) usability challenges and teething issues can exacerbate stress for first-time researchers. The lead author's experience highlights the importance of adequate training and support, particularly for HDR students balancing research, employment and personal responsibilities. These reflections extend the literature on ethics administration by highlighting how digital platforms can unintentionally create barriers that impact research timelines and student wellbeing.

Social Media Participant Recruitment

The use of social media for participant recruitment illustrates the potential and the ethical complexities of digital platforms in research. Social media allows access to a diverse range of the population and facilitates rapid, low-cost recruitment (Benedict et al., 2019; Darko et al., 2022; Gelinás et al., 2017). However, the lead author's experience reveals how posts can attract participants outside the intended demographic, raising questions of consent, privacy and data inclusion. This aligns with research lighting that digital recruitment can blur the boundaries between personal and professional spaces, complicating informed consent and engagement with participants in an authentic way (Gelinás et al., 2017; Zimmermann et al., 2022). By revising the recruitment post to explicitly target Australian educators and consulting with ethical guidelines (Association of Internet Researchers), the research team demonstrated reflexivity in addressing these micro-ethical challenges. These actions provide a model for other researchers in navigating the ethical tensions inherent in online recruitment,

particularly within sensitive professional contexts like early childhood education and care.

Geo-Location Data and Inclusion Criteria

The discovery of the geo-location data embedded within survey responses by participants presented a micro-ethical dilemma regarding inclusion and exclusion. Although the survey was intended for Australian educators, responses from international participants required filtering, balancing the respect for participants' time and contribution with the integrity of the research aims. Literature on ethical data management emphasises the importance of clear inclusion/exclusion criteria and reflexivity in handling unexpected data (Burke et al., 2024; Taquette & Borges da Matta Souza, 2022). This situation demonstrated how micro-ethical dilemmas emerge during data collection, necessitating ongoing dialogue between the HDR student and the supervisory team to ensure decisions are ethically justified, documented and consistent with the study's ethics approval and objectives.

Reflexivity, Micro-Ethical Dilemmas, and Research Integrity

Across the three reflections, the concept of micro-ethical dilemmas is central. These subtle and often unanticipated challenges require researchers to engage in ethical reflexivity, considering what is allowed under formal ethical guidelines, and also what is respectful, contextually sensitive and aligned with professional norms (Burke et al., 2024). By documenting these dilemmas, the lead author contributes to methodological knowledge by demonstrating how HDR students can navigate ethical complexities, where ethical choices involve balancing transparency, anonymity and the reputational consideration of participants and institutions.

Implications for the ECEC Sector and Future Research

These reflections have some implications for researchers working within the ECEC sector, particularly those engaging in digitally mediated research and social media recruitment. First, the reflections highlight the importance of anticipating ethical complexity beyond procedural approval processes. Researchers need to plan for micro-ethical decision-making throughout the cycle of research, including recruitment and dissemination. This is particularly relevant in online environments where issues such as privacy, consent and geo-location can emerge unexpectedly during data collection.

Secondly, the findings emphasise the need for more intentional and transparent recruitment planning when using social media. Researchers should carefully consider how platforms can blur the line between professional and personal, and explicitly design

recruitment materials that minimise coercion, protect anonymity and separate the researcher's role from informal online interactions. This includes sharing consent processes that are layered and revisited, using plain language in the participant information sheet and building clear opt-out mechanisms at multiple stages of participation within the online survey.

Finally, the reflections suggest a need for stronger methodological preparation for early career researchers and higher degree research students conducting digital research. Rather than a sole focus on procedural ethics compliance, research training programs may look at including discussions of ethics in practice, including case exploration of micro-ethical dilemmas, and digital consent. Additionally, supervisory relationships play a crucial role in supporting reflexive ethical judgement, particularly when encountering unanticipated challenges during data collection and dissemination.

Methodological and ethical decision making are inseparable from research design and implementation. Ethical practice emerges continuously throughout the process rather than being confined to the formal approval stages in the beginning of research and therefore requires ongoing reflexivity and adaptive decision-making. These insights offer a practical guide for emerging and established researchers seeking to conduct ethically responsible and methodologically rigorous research within digital environment in the ECEC context.

AI Disclosure Statement: Artificial Intelligence (AI) tools (ChatGPT) were used to assist with language refinement and grammar. All intellectual content, study design, analysis, interpretation of findings and final editorial decisions were undertaken solely by the authors.

References

- Andrew, Y. (2015). What we feel and what we do: Emotional capital in early childhood work. *Early Years*, 35(4), 351-365. <https://doi.org/10.1080/09575146.2015.1077206>
- Australian Children's Education and Care Quality Authority (ACECQA). (2019). *Progressing a national approach to the children's education and care workforce*. A. Government. <https://www.acecqa.gov.au/sites/default/files/2022-01/ProgressingNationalApproachChildrensEducationCareWorkforce.pdf>
- Benedict, C., Hahn, A. L., Diefenbach, M. A., & Ford, J. S. (2019). Recruitment via social media: advantages and potential biases. *Digital health*, 5, 2055207619867223. <https://journals.sagepub.com/doi/epub/10.1177/2055207619867223>

- Berger, R. (2015). Now I see it, now I don't: Researcher's position and reflexivity in qualitative research. *Qualitative research*, 15(2), 219-234. https://www.researchgate.net/profile/Roni-Berger/publication/260984361_Now_I_see_it_now_I_don%27t_researcher%27s_position_and_reflexivity_in_qualitative_Research/links/00b4953734f1ab1244000000/Now-I-see-it-now-I-dont-researchers-position-and-reflexivity-in-qualitative-Research.pdf
- Bertram, T., Pascal, C., Lyndon, H., Formosinho, J., Gaywood, D., Gray, C., Koutoulas, J., Loizou, E., Vandenbroek, M., & Whalley, M. (2025). EECERA ethical code for early childhood researchers. *European Early Childhood Education Research Journal*, 33(1), 4-18. <https://doi.org/10.1080/1350293X.2024.2445361>
- Buchanan, E. A., & Zimmer, M. (2012). Internet research ethics. <https://plato.stanford.edu/ENTRIES/ethics-internet-research/>
- Bull, R., McFarland, L., Cumming, T., & Wong, S. (2024). The impact of work-related wellbeing and workplace culture and climate on intention to leave in the early childhood sector. *Early Childhood Research Quarterly*, 69, 13-24. <https://doi.org/10.1016/j.ecresq.2024.06.002>
- Burke, R., Baker, S., Molla, T., Cabiles, B., & Fox, A. (2024). How do higher degree research students and supervisors navigate ethics-in-practice for educational research in sensitive or 'fragile' contexts? *British Educational Research Journal*, 50(2), 837-854. <https://bera-journals.onlinelibrary.wiley.com/doi/full/10.1002/berj.3945>
- Darko, E. M., Kleib, M., & Olson, J. (2022). Social media use for research participant recruitment: integrative literature review. *Journal of Medical Internet Research*, 24(8), e38015. <https://www.jmir.org/2022/8/e38015/PDF>
- Doyle, L., Brady, A-M., & Byrne, G. . (2009). An overview of mixed methods research *Journal of Research in Nursing* 14(2), 175-185. <https://doi.org/10.1177/174498718093962>
- Early Childhood Australia. (2022). *Digital Child Ethics Toolkit: Ethical Considerations for Digital Childhoods Research*. <https://thespokeearlychildhoodaustralia.org.au/digital-child-ethics-toolkit-considerations-for-research>
- Finlay, L. (2002). "Outing" the researcher: The provenance, process, and practice of reflexivity. *Qualitative health research*, 12(4), 531-545. <https://doi.org/10.1177/104973202129120052>
- Franzke, A. S., Bechmann, A., Ess, C. M., & Zimmer, M. (Eds.) (2020). Internet Research: Ethical Guidelines 3.0. AoIR (The International Association of Internet Researchers). <https://aoir.org/reports/ethics3.pdf>
- Gelinas, L., Pierce, R., Winkler, S., Cohen, I. G., Lynch, H. F., & Bierer, B. E. (2017). Using social media as a research recruitment tool: ethical issues and recommendations. *The American journal of bioethics*, 17(3), 3-14. <https://pmc.ncbi.nlm.nih.gov/articles/PMC5324729/pdf/nihms845256.pdf>
- Guillemin, M., & Gillam, L. (2004). Ethics, reflexivity, and "ethically important moments" in research. *Qualitative Inquiry*, 10(2), 261-280. <https://journals.sagepub.com/doi/epdf/10.1177/1077800403262360>
- Kapoor, K. K., Tamilmani, K., Rana, N. P., Patil, P., Dwivedi, Y. K., & Nerur, S. (2018). Advances in social media research: Past, present and future. *Information systems frontiers*, 20, 531-558. <https://link.springer.com/content/pdf/10.1007/s10796-017-9810-y.pdf>
- Lyndon, H. (2023). Embracing the breadth of ethical complexities in early childhood research. In (Vol. 31, pp. 143-146): Taylor & Francis.
- Markham, A., & Buchanan, E. (2012). Ethical decision-making and internet research: Recommendations from the AoIR Ethics Working Committee (Version 2.0). <https://pure.au.dk/ws/files/55543125/aoirethics2.pdf>
- McMullen, M. B., Lee, M. S., McCormick, K. I., & Choi, J. (2020). Early childhood professional well-being as a predictor of the risk of turnover in child care: A matter of quality. *Journal of Research in Childhood Education*, 34(3), 331-345. https://www.researchgate.net/profile/Mary-Mcmullen/publication/338956362_Early_Childhood_Professional_Well-Being_as_a_Predictor_of_the_Risk_of_Turnover_in_Child_Care_A_Matter_of_Quality/links/5fc148eda6dcc6cc676530b/Early-Childhood-Professional-Well-Being-as-a-Predictor-of-the-Risk-of-Turnover-in-Child-Care-A-Matter-of-Quality.pdf
- Menning, S. F., Caetano-Silva, G., & Murriss, K. (2025). 'What action did you see me do?': proposing a methodology of choreographed visual montage when re-turning to interview research data. *European Early Childhood Education Research Journal*, 1-19. <https://www.tandfonline.com/doi/full/10.1080/1350293X.2025.2502821>

- Mérida-López, S., & Extremera, N. (2017). Emotional intelligence and teacher burnout: A systematic review. *International Journal of Educational Research*, 85, 121-130. <https://www.sciencedirect.com/science/article/abs/pii/S0883035517304214?via%3Dihub>
- Rani, M., Chawla, N., Wadhwa, N., Mathur, R., Jinks, T., Das, P., & Rijal, S. (2024). A transformative solution to build effective, transparent, and resilient “fit-for-purpose” national health research ethics systems. *Health Research Policy and Systems*, 22(1), 131. <https://link.springer.com/article/10.1186/s12961-024-01219-2>
- Sheehy, A., James, J. R., & Horgan, M. (2020). Implementing a national approach to research ethics review during a pandemic—the Irish experience. *HRB Open Research*, 3, 63. <https://pmc.ncbi.nlm.nih.gov/articles/PMC7797935/>
- Taquette, S. R., & Borges da Matta Souza, L. M. (2022). Ethical Dilemmas in Qualitative Research: A Critical Literature Review. *International Journal of Qualitative Methods*, 21, 16094069221078731. <https://doi.org/10.1177/16094069221078731>
- Van De Mierop, D., Clifton, J., & Schnurr, S. (2022). Narratives as social practice in organisational contexts. *Narrative Inquiry*, 32(1), 1-8. https://www.jbe-platform.com/docserver/fulltext/ni.21090.van.pdf?exp_ires=1652907835&id=id&accname=g_uest&checksum=412E94194395ACA0DED167F0A7D5E85B
- Whitaker, C., Stevelink, S., & Fear, N. (2017). The use of Facebook in recruiting participants for health research purposes: a systematic review. *Journal of Medical Internet Research*, 19(8), e290. <https://www.jmir.org/2017/8/e290/PDF>
- Zimmermann, B. M., Willem, T., Bredthauer, C. J., & Buyx, A. (2022). Ethical issues in social media recruitment for clinical studies: ethical analysis and framework. *Journal of Medical Internet Research*, 24(5), e31231. <https://www.jmir.org/2022/5/e31231/>

Parent Engagement in Early Childhood Education: A Multivariate Analysis in the Turkish Context*

Ümit Ünsal Kaya^a, Ümmühan Akpınar Afşin^{b,**}

Received : 17 January 2026
Revised : 27 June 2026
Accepted : 20 July 2026
DOI : 10.26822/iejee.2026.465

^aÜmit Ünsal Kaya, School of Foreign Languages,
Afyon Kocatepe University, Afyonkarahisar, Türkiye.
E-mail: uakpinar@bartin.edu.tr
ORCID: <https://orcid.org/0000-0001-7662-8089>

^{b,**} **Corresponding Author:** Ümmühan Akpınar Afşin,
Department of Early Childhood Education, Bartın
University, Bartın, Türkiye.
E-mail: ummuhanakpinar.ua@gmail.com
ORCID: <https://orcid.org/0000-0001-8107-6004>

* This study, under the same title, was presented
as an oral paper at the 9th International Congress
on Preschool Education, hosted by the Faculty
of Education at Bursa Uludağ University, held on
October 27–30, 2025.

Abstract

This study examined parent engagement in early childhood education in Türkiye and its associations with demographic characteristics. A cross-sectional survey was conducted with 496 parents of children aged 66–72 months recruited through convenience sampling with snowball features. Participants completed the 25-item Turkish Parent Engagement in Early Childhood Education Scale, comprising Knowledge and Expectations, Trust and Communication, and Home-Based Engagement. Demographic differences were examined using Mann–Whitney U and Kruskal–Wallis tests with false discovery rate correction, followed by multiple linear regression analyses with HC3 standard errors. Overall engagement was high, although Home-Based Engagement was comparatively lower. Respondents with one child reported higher Home-Based Engagement than respondents with two, three, four or more children, whereas the multiple-child groups did not differ significantly from one another. Urban respondents also reported higher Home-Based Engagement than rural respondents. Educational-level contrasts were identified for Total Parent Engagement but not for Home-Based Engagement after adjustment. Home-Based Engagement demonstrated low internal consistency ($\alpha = .48$; $\omega = .52$). Therefore, findings involving this subdimension should be considered preliminary. The results indicate demographic associations rather than causal effects and require replication using psychometrically stronger measures and more representative samples.

Keywords:

Early Childhood Education; Demographic Variables; Parental Engagement; Parent Involvement; School-Family Relationship

Introduction

Early childhood education is widely recognized as a critical period for children's cognitive, social, and emotional development, during which learning experiences are shaped not only by educational institutions but also by family contexts. Within this period, parents play a central role in supporting children's learning trajectories by shaping everyday experiences, learning opportunities, and relationships across home and school settings. Accordingly, parent engagement has been consistently identified as a key mechanism through which early educational experiences are enriched and sustained (Fantuzzo et al., 2013; Weiss et al., 2010). Rather than being limited to



www.iejee.com
ISSN: 1307-9298

2026 Published by KURA Education & Publishing.
This is an open access article under the CC BY-
NC- ND license. (<https://creativecommons.org/licenses/by/4.0/>)

school-based participation, parent engagement encompasses a broader set of beliefs, practices, and relational processes that influence children's development over time and across multiple contexts (Manzon et al., 2015).

Despite its recognized importance, parent engagement remains conceptually diverse and is sometimes used interchangeably with family engagement and parent involvement (Fantuzzo et al., 2013; Gross et al., 2020; Halgunseth et al., 2013). In the present study, parent engagement is understood as the beliefs, attitudes, relationships, and activities through which parents support children's learning across home and educational settings. This conceptualization emphasizes sustained, bidirectional, and trust-based relationships between families and educational programs rather than participation in isolated school-defined activities (Manzon et al., 2015; Weiss et al., 2010). It therefore provides a multidimensional framework encompassing parents' knowledge and expectations, communication with educational institutions, and learning-related practices at home.

Theoretical frameworks further contribute to understanding how parent engagement is established and sustained in early childhood contexts. Two influential approaches have been identified in the literature: Weiss and colleagues' framework (2010), which primarily addresses elementary and secondary education, and Halgunseth and colleagues' Social Exchange Model of Family Engagement (2009), which is specifically grounded in early childhood education settings (Forry et al., 2012). Given the focus of the present study, the Social Exchange Model of Family Engagement provides a particularly relevant lens. Drawing on social exchange and ecological theories, this model conceptualizes parent engagement as emerging from reciprocal, trust-based partnerships between families and early childhood programs, supported by both program-level and family-level resources (Halgunseth et al., 2009). According to the model, engagement is fostered through practices such as welcoming school environments, two-way communication, home visits, joint decision-making, parent education opportunities, and the provision of home-based learning resources, while families contribute by supporting learning at home, sharing information with educators, participating in program activities, and collaborating with schools (Halgunseth et al., 2009). Importantly, the model highlights the need to attend to cultural sensitivity and contextual diversity, suggesting that parent engagement may vary meaningfully across demographic and environmental conditions. Recent evidence from low-resource settings similarly suggests that parental engagement is patterned by socioeconomic conditions. For example, using an ecological framing and multivariate logistic regression, Gitonga et al. (2026) examined parents of Pre-Primary 2 children

in public ECE settings in rural Kenya and found that higher parental education and household income were associated with higher engagement, whereas economic constraints substantially limited parents' ability to contribute to their children's early education. Such findings underscore the relevance of examining demographic and structural correlates of engagement in a country-specific manner.

Empirical findings also suggest that family engagement may vary according to parental and household characteristics, although the direction of these associations is not always consistent across contexts or engagement dimensions. In Türkiye, Varol et al. (2023) found that mothers reported greater participation than fathers and that maternal education was associated with school-based involvement, with primary-school-educated mothers reporting greater participation than mothers with bachelor's degrees. Other Turkish studies have identified employment demands, limited time, communication difficulties, and competing family responsibilities as barriers to sustained participation (Çakmak, 2010; Orçan Kaçan et al., 2019; Akyol et al., 2021). Recent qualitative evidence from rural Türkiye has additionally highlighted spatial inequalities in access, adequacy, and educational opportunities experienced by families (Dikmen & Gökalp, 2026). Taken together, these findings suggest that educational background, family responsibilities, and residential context may shape both families' opportunities for engagement and the forms that engagement takes. However, these relationships remain insufficiently examined through multidimensional measures of parent engagement in Turkish early childhood education.

Consistent with these theoretical perspectives, empirical research has conceptualized parent engagement as a multidimensional construct encompassing various forms of interaction across home, school, and community contexts. For instance, Gross et al. (2022) synthesized the literature and identified core dimensions of parent engagement, including home-based engagement, school-based engagement, and communication among parents, teachers, and schools. Home-based engagement refers to parents' efforts to support learning within the home environment, whereas school-based engagement involves participation in activities organized by educational institutions (Gross et al., 2022; Wondim, 2024). Similarly, the OECD (2011), building on earlier work by Epstein (1995) and Oakes and Lipton (2007), distinguished between child-focused practices such as parenting, communicating, and fostering learning at home, and center-oriented practices such as volunteering, decision-making, and collaboration with the community. Other studies have emphasized specific manifestations of engagement, including volunteering, parent-teacher meetings, classroom activities, and home-based learning support (Cutshaw

et al., 2022; Wondim, 2024). Taken together, these perspectives underscore that parent engagement is not a singular behavior but a constellation of practices that may differ by context, educational level, and stakeholder perspective (OECD, 2011).

Parent engagement should also be distinguished from parent involvement. Although the terms are sometimes used interchangeably, involvement generally refers to parents' participation in school-defined activities, whereas engagement emphasizes shared responsibility, parental agency, and sustained collaboration around children's learning (Ferlazzo, 2011; Goodall & Montgomery, 2014). Parent engagement therefore extends beyond attendance or compliance and foregrounds reciprocal relationships between families and educational professionals (Gross et al., 2022). Consistent with this distinction, the present study focuses on parent engagement.

In Türkiye, parent participation in education has gradually developed through policy, institutional, and community-based initiatives. School–family associations, parent meetings, home visits, parent education programs, and the work of organizations such as the Mother–Child Education Foundation have contributed to this development (Kağıtçıbaşı et al., 2009; Tekin, 2011). Successive early childhood education programs have also moved from treating parent involvement as a general principle toward providing more structured guidance for family and community participation (MoNE, 2002; 2013; 2024a; 2024b). Nevertheless, parent involvement remains the dominant term in Turkish policy and research, and multidimensional parent engagement has received comparatively limited empirical attention (Kaya, 2023; Kaya, 2024; Serdaroğlu & Bekir, 2024).

Against this backdrop, there is a notable gap in the empirical literature concerning parent engagement in early childhood education in Türkiye. To the best of the authors' knowledge, few studies have systematically examined parent engagement as a multidimensional construct across demographic characteristics using a validated measurement framework in the Turkish early childhood education context. Existing studies have often focused on parent involvement, relied on limited indicators, or lacked comprehensive demographic analyses. Consequently, little is known about how different dimensions of parent engagement, particularly home-based engagement, vary across family characteristics such as parental education, family size, and residential context. Addressing this gap is essential for advancing the international literature on parent engagement and for informing contextually responsive policies and practices in early childhood education in Türkiye.

The present study aims to address this need by examining parental engagement in early childhood education within the Turkish context using a

multidimensional, validated scale. By investigating overall levels of parent engagement and exploring demographic differences across dimensions, the study contributes to the literature by providing a detailed demographic profile of multidimensional parent engagement in early childhood education in Türkiye. In addition to its theoretical contribution, the study offers practical implications for early childhood programs and policymakers by identifying demographic patterns associated with home-based engagement, thereby highlighting groups that may benefit from targeted support and intervention. Guided by this rationale, the study addresses the following research questions: (1) Do Total Parent Engagement and subdimension scores differ across demographic characteristics in the Turkish context? (2) Which demographic variables are associated with Home-Based Engagement? (3) Which demographic characteristics show adjusted associations with parent engagement outcomes when considered simultaneously?

Method

Research Design

This study employed a quantitative cross-sectional correlational survey design to describe parent engagement and examine its associations with selected demographic characteristics. Data were collected at a single point in time without manipulating the study variables. Accordingly, the findings represent descriptive and relational patterns and do not support causal inference (Creswell & Creswell, 2018; Cohen et al., 2018). Parent engagement was examined as a multidimensional construct encompassing Knowledge and Expectations, Trust and Communication, and Home-Based Engagement. Demographic group differences were evaluated using nonparametric procedures, and adjusted associations were examined using multivariable regression models.

Participants

The participants of the study consisted of 496 parents of children aged 66 to 72 months residing in different regions of Türkiye. Participants were recruited using a convenience sampling strategy. Because the survey link was shared among participants' personal and social networks, the recruitment process also exhibited characteristics of snowball sampling. This approach increased accessibility and facilitated data collection across a wide geographic area; however, it also imposed limitations on the representativeness of the sample and, consequently, on the generalizability of the findings. Data collection was conducted online in April 2025. Participation was entirely voluntary, and no incentives were provided. All participants were informed about the purpose of the study and completed the questionnaire after providing informed consent. Detailed demographic characteristics of

the participants are presented in Table 1. These characteristics include parent gender, monthly household income, parent age, number of children in the household, parent employment status, parent educational level, child gender, prior preschool education experience, type of preschool institution, and place of residence, categorized as urban or rural.

Table 1.
Participants' Demographic Characteristics (N = 496)

Variable	Category	n	%
Parent Gender	Male	45	9.1
	Female	451	90.9
Monthly Household Income	Low	249	50.2
	Middle	188	37.9
	High	59	11.9
Parent Age (years)	22–25	29	5.8
	26–30	126	25.4
	31–35	200	40.3
	36–40	103	20.8
	41–45	35	7.1
	46 and above	3	0.6
Number of Children in the Household	1	139	28.0
	2	247	49.8
	3	85	17.1
	4 or more	25	5.0
Parent Employment Status	Public sector employee / Worker	90	18.1
	Not employed / Home-maker	307	61.9
	Private sector employee	99	20.0
Parent Educational Level	Doctorate	5	1.0
	High school	148	29.8
	Middle school	77	15.5
	Master's degree	17	3.4
	Bachelor's degree	205	41.3
	Primary school	44	8.9
Child Gender	Male	271	54.6
	Female	225	45.4
Prior Preschool Education Experience	Yes	220	44.4
	No	276	55.6
Type of Preschool Institution	Public	367	74.0
	Private	129	26.0
Place of Residence	Rural (town/village)	75	15.1
	Urban (city center)	421	84.9

An examination of Table 1 indicates that the majority of participants were mothers, nearly half of the households were in the lower income category, and the most common age group among parents was 31 to 35 years. More specifically, mothers constituted 90.9% of the sample. The term “parents” is retained because both mothers and fathers participated; however, the findings primarily reflect maternal reports. To examine whether the principal regression patterns were

driven by the small group of fathers, the multivariable analyses were repeated in the mother-only subsample as a sensitivity analysis. Approximately half of the families had two children, and most participants resided in urban areas. While this demographic distribution reflects several typical characteristics of parents of preschool-aged children in Türkiye, the use of convenience sampling with snowball features requires that the findings be interpreted with caution, particularly with regard to their generalizability to the broader population. Because the master's and doctoral groups were small, these categories were combined with the bachelor's group as “bachelor's degree and above” for inferential analyses. For the regression analyses, the six descriptive parent-age categories presented in Table 1 were consolidated into three broader groups: 22–30 years (n = 155), 31–35 years (n = 200), and 36 years and above (n = 141). This consolidation avoided the unstable estimation that would result from retaining the very small oldest age category.

Data Collection Tool

The data collection tool used in this study was the Early Childhood Education Parent Engagement Scale, referred to in Turkish as the Erken Çocukluk Eğitime Aile Bağlılığı Ölçeği. This scale is the Turkish adaptation of the Parent Engagement in Early Childhood Education Scale originally developed by Gross and colleagues (Gross et al., 2020; Gross et al., 2022). The Turkish adaptation, validity, and reliability study of the scale was conducted by Kaya, Demircan, and Altun (2024). As a result of the adaptation process, the Turkish version of the scale retained the original three-dimensional structure and was culturally adapted and validated for use in the Turkish context.

The Early Childhood Education Parent Engagement Scale is designed to assess parents' levels of engagement in early childhood education, with a particular focus on their involvement in their children's learning processes, school–family collaboration, and home-based educational activities. The scale consists of a total of twenty-five items, and responses are obtained using a four-point Likert-type format ranging from one, strongly disagree, to four, strongly agree. Higher scores indicate higher levels of parental engagement in early childhood education.

The scale comprises three subdimensions. The Knowledge and Expectations subdimension consists of nine items and measures parents' awareness and knowledge regarding their child's learning process, educational goals, and expectations from the school. The Trust and Communication subdimension consists of eleven items and assesses the level of trust-based communication parents establish with teachers and school administrators. The Home-Based Engagement subdimension consists of five items and captures

parents' behaviors related to supporting their child's learning at home, engaging in interactions with their child, and creating a learning-supportive home environment.

Evidence for the validity and reliability of the Turkish version of the scale was provided in the adaptation study conducted by Kaya et al. (2024), in which confirmatory factor analysis supported the three-factor structure of the instrument, with acceptable model fit indices reported as a chi-square to degrees of freedom ratio equal to 2.84, a comparative fit index equal to .93, and a root mean square error of approximation equal to .057. In the same study, internal consistency coefficients were reported as .82 for the total scale and ranged between .70 and .80 for the subdimensions.

The psychometric characteristics of the scores were re-examined in the present sample. Internal consistency was assessed using raw Cronbach's alpha and McDonald's omega. In addition, corrected item-total correlations, mean inter-item correlations, and alpha-if-item-deleted coefficients were examined. Because Home-Based Engagement demonstrated low internal consistency, item-level response distributions and the previously established three-factor structure were also evaluated. The complete sample-specific psychometric findings are presented in the Results section.

Data Collection Process

Data collection was carried out online in April 2025 using a digital survey platform (Google Forms). The survey link was distributed via social media platforms, online parent groups, and communication networks commonly used by parents of preschool-aged children. Before accessing the questionnaire, participants were presented with an informed consent form that explained the purpose of the study, voluntary nature of participation, confidentiality of responses, and the right to withdraw at any time without penalty. Only participants who provided consent were able to proceed to the survey. No personally identifiable information was collected. All responses were recorded anonymously and used solely for scientific research purposes. The average completion time for the questionnaire was approximately 10–12 minutes. Ethical approval for the study was obtained from the Afyon Kocatepe University Social and Human Sciences Scientific Research and Publication Ethics Committee (Approval No: 2024/70). The study was conducted in accordance with ethical principles governing research with human participants.

Data Analysis

Prior to the analyses, the dataset was screened, coded, and checked for completeness. For inferential

analyses, participants with bachelor's, master's, and doctoral degrees were combined into a "bachelor's degree and above" category to avoid unstable estimates based on sparse postgraduate groups. The resulting educational-level variable comprised primary school, middle school, high school, and bachelor's degree and above. The six descriptive parent-age categories presented in Table 1 were also consolidated into three broader groups for the regression analyses: 22–30 years, 31–35 years, and 36 years and above. This recategorization eliminated the sparsely populated oldest age category and provided more stable estimates. All categorical predictors were dummy-coded. The regression models therefore included one parameter for parent gender, two for household income, two for parent age, two for employment status, three for educational level, three for number of children, and one each for child gender, prior preschool education experience, institution type, and place of residence, yielding 17 predictor parameters in total. The model intercept was estimated separately and was not included in the numerator degrees of freedom.

In the first stage, descriptive statistics were computed for the total scale score and each subdimension, including means, standard deviations, medians, minimum and maximum values, and skewness and kurtosis coefficients. Distributional assumptions were evaluated using the Shapiro–Wilk test (Shapiro & Wilk, 1965). The results indicated significant deviations from normality for all scale scores ($p < .001$). In addition, the distributions were negatively skewed and showed clustering toward the upper end of the response scale, suggesting potential ceiling effects. These distributional characteristics were taken into account when selecting subsequent inferential procedures (Field, 2013).

Internal consistency was evaluated using raw Cronbach's alpha and McDonald's omega. Raw alpha coefficients were calculated from the item covariance matrix, whereas omega coefficients were estimated using Pearson item correlations. Corrected item-total correlations and mean inter-item correlations were also based on Pearson correlations. Alpha-if-item-deleted coefficients and item-level response distributions were examined for Home-Based Engagement. Because raw alpha incorporates differences in item variances, it was not expected to correspond exactly to the standardized alpha implied by the mean inter-item correlation. An ordinal confirmatory factor analysis was conducted using the weighted least squares mean- and variance-adjusted estimator to examine the previously established three-factor structure. Model fit was evaluated using the comparative fit index, Tucker–Lewis index, root mean square error of approximation with its 90% confidence interval, and standardized root mean square residual. Standardized

factor loadings were also examined, with particular attention to the Home-Based Engagement items.

Demographic group differences were examined using the Mann–Whitney U test for two-group comparisons and the Kruskal–Wallis H test for comparisons involving three or more groups. The Benjamini–Hochberg false discovery rate correction was applied separately to the 10 demographic comparisons conducted for each engagement outcome. FDR-adjusted demographic group-comparison results were reported as q values. When a Kruskal–Wallis test remained significant after FDR correction, Dunn pairwise comparisons were conducted, and the resulting p values were separately adjusted using the Benjamini–Hochberg procedure within the relevant set of post hoc comparisons. Effect sizes were reported as r for Mann–Whitney U tests and ordinal epsilon-squared for Kruskal–Wallis tests. Ordinal epsilon-squared was calculated as $\epsilon^2 = H/(N - 1)$.

Multiple linear regression models were estimated using ordinary least squares to examine adjusted associations between demographic characteristics and the parent engagement outcomes. Separate models were estimated for Total Parent Engagement and each subdimension. Conventional overall OLS F tests were used to evaluate model-level statistical significance. Because the outcomes were bounded averaged Likert-type scores and showed ceiling effects, HC3 heteroskedasticity-consistent standard errors were used for coefficient-level inference. Unstandardized coefficients, HC3 standard errors, 95% confidence intervals, and coefficient-level p values were reported together with R^2 and adjusted R^2 . Residual-versus-fitted plots and Cook's distance values were examined as diagnostic checks. Statistical interpretations were based on individual coefficients and reference-group contrasts rather than unreported omnibus tests for multi-category predictors. Because mothers constituted 90.9% of the sample, the multivariable regression analyses were repeated using only the 451 maternal respondents. This sensitivity analysis was conducted to determine whether the principal regression patterns were influenced by the small group of fathers. As in the primary analyses, HC3 standard errors were used for coefficient-level inference.

Results

Preliminary Analyses

Preliminary analyses were conducted to examine the distributional and psychometric characteristics of the Parent Engagement in Early Childhood Education Scale scores. As presented in Table 2, the mean Total Parent Engagement score was 3.73 (SD = 0.21). The mean subdimension scores were 3.74 (SD = 0.25) for Knowledge and Expectations, 3.77 (SD = 0.28) for Trust

and Communication, and 3.63 (SD = 0.29) for Home-Based Engagement. Home-Based Engagement therefore had the lowest mean score, although scores across all dimensions were concentrated near the upper end of the four-point scale.

The Shapiro–Wilk tests indicated significant departures from normality for the total score and all three subdimensions (all $p < .001$). The distributions were negatively skewed, particularly for Total Parent Engagement, Knowledge and Expectations, and Trust and Communication. Given these distributional characteristics and the observed ceiling effects, nonparametric tests were used for demographic group comparisons.

Internal Consistency and Supplementary Psychometric Analyses

Table 3 presents the internal consistency estimates for the scale scores. Cronbach's alpha was .78 for the total scale, .61 for Knowledge and Expectations, .77 for Trust and Communication, and .48 for Home-Based Engagement. The corresponding McDonald's omega estimates were .80, .64, .79, and .52. Although omega produced slightly higher estimates than alpha, it did not eliminate the reliability concern associated with Home-Based Engagement.

Because Home-Based Engagement demonstrated low internal consistency, additional item-level analyses were conducted. Corrected item–total correlations for its five items ranged from .16 to .38, and the mean Pearson inter-item correlation was .17. Alpha-if-item-deleted coefficients ranged from .39 to .51. Removing the item with the weakest corrected item–total correlation increased alpha only marginally, from .48 to .51. No item was therefore removed because the improvement was limited and deleting the item would have altered the content coverage of the original validated subdimension.

Item-level distributions indicated substantial clustering in the highest response category. Depending on the item, between 61.5% and 82.7% of respondents selected the highest response option. This restricted variability may have contributed to the weak inter-item correlations, although it cannot fully account for the low reliability estimate.

An ordinal confirmatory factor analysis using the weighted least squares mean- and variance-adjusted estimator was conducted to examine the previously established three-factor structure. The model demonstrated acceptable overall fit, $\chi^2(272) = 451.86$, $p < .001$, CFI = .955, TLI = .950, RMSEA = .037, 90% CI [.031, .043], and SRMR = .051. Standardized factor loadings ranged from .53 to .81 for Knowledge and Expectations, from .49 to .79 for Trust and Communication, and from .31 to .59 for Home-Based Engagement.

The CFA supported the overall three-factor structure. However, the comparatively weak factor loadings and low internal consistency of Home-Based Engagement indicated that its composite score contained considerable measurement error. Findings involving this subdimension were therefore regarded as preliminary and interpreted cautiously.

Demographic Differences in Parent Engagement

To obtain more stable estimates, participants with bachelor's, master's, and doctoral degrees were combined into a "bachelor's degree and above" category. The revised educational-level variable comprised primary school ($n = 44$), middle school ($n = 77$), high school ($n = 148$), and bachelor's degree and above ($n = 227$).

Following this recategorization, Total Parent Engagement differed significantly across educational groups, $H(3) = 14.82$, $p = .002$, FDR-adjusted $q = .021$, $\epsilon^2 = .030$. Although statistically significant, the effect size was small. FDR-adjusted Dunn comparisons indicated that respondents with bachelor's degrees or above reported lower Total Parent Engagement than respondents with primary school education (adjusted $p = .009$). The remaining pairwise differences were not statistically significant after correction.

Home-Based Engagement differed significantly according to the number of children in the household, $H(3) = 37.21$, $p < .001$, $q < .001$, $\epsilon^2 = .075$. This was the largest demographic effect identified in the group comparisons. FDR-adjusted Dunn tests showed that respondents with one child reported higher scores than respondents with two children (adjusted $p = .006$), three children (adjusted $p < .001$), and four or more children (adjusted $p = .003$). No statistically significant differences were found among the two-child, three-child, and four-or-more-child groups after correction. The findings therefore indicated a distinction between one-child and multiple-child households rather than a strictly monotonic decline across all categories.

Place of residence was also associated with Home-Based Engagement. Respondents living in urban areas reported higher scores than those living in rural areas, $U = 12,554.50$, $p = .004$, $q = .013$, $r = .12$. The magnitude of this difference was small.

Home-Based Engagement also differed across the revised educational groups, $H(3) = 17.96$, $p < .001$, $q = .002$, $\epsilon^2 = .036$. FDR-adjusted Dunn comparisons indicated that respondents with bachelor's degrees or above reported higher Home-Based Engagement than respondents with middle school education (adjusted $p = .018$) and high school education (adjusted $p = .011$). The difference between the primary school and bachelor's-degree-and-above groups was not statistically significant after correction (adjusted $p =$

.083). Given the limited reliability of the subdimension, these educational differences should be interpreted cautiously.

No statistically significant differences remained after FDR correction for parent gender, child gender, household income, parent age, employment status, prior preschool education experience, or institution type. Similarly, no FDR-adjusted demographic differences were identified for Knowledge and Expectations or Trust and Communication.

Demographic Variables Associated With Home-Based Engagement

The second research question focused specifically on the demographic variables associated with Home-Based Engagement. The clearest difference concerned the number of children in the household. Respondents with one child reported higher Home-Based Engagement than respondents with multiple children. However, the multiple-child groups did not differ significantly from one another after correction.

Urban respondents also reported higher Home-Based Engagement than rural respondents, although this difference was small. Educational level was associated with Home-Based Engagement at the univariate level, with the bachelor's-degree-and-above group reporting higher scores than the middle school and high school groups.

These findings indicate demographic associations rather than causal effects. Moreover, because Home-Based Engagement demonstrated low internal consistency, the magnitude and stability of these differences remain uncertain.

Multivariable Analyses

Multiple linear regression models were estimated using ordinary least squares to examine the adjusted associations between demographic characteristics and parent engagement outcomes. Conventional overall OLS F tests were used to evaluate model-level statistical significance. HC3 heteroskedasticity-consistent standard errors were used for coefficient-level inference, including standard errors, confidence intervals, and p values. All demographic variables were entered simultaneously, and the revised four-category educational-level variable was used in each model. The numerator degrees of freedom correspond to the 17 predictor parameters entered into each model after category consolidation and dummy coding; the intercept was not included in this count.

Total Parent Engagement

The overall OLS model predicting Total Parent Engagement was statistically significant, $F(17,$

478) = 4.03, $p < .001$, $R^2 = .125$, adjusted $R^2 = .094$. The coefficient-level results reported below are based on HC3 heteroskedasticity-consistent standard errors.

Compared with respondents who had completed primary school, respondents with middle school education reported lower Total Parent Engagement, $b = -0.146$, robust SE = 0.045, 95% CI [-0.235, -0.057], $p = .001$. Lower scores were also observed among respondents with high school education, $b = -0.166$, robust SE = 0.044, 95% CI [-0.253, -0.079], $p < .001$, and those with bachelor's degrees or above, $b = -0.153$, robust SE = 0.043, 95% CI [-0.238, -0.068], $p < .001$.

No other individual demographic coefficient or reference-group contrast was statistically significant. These coefficients should be interpreted cautiously because the model explained only a modest proportion of the variance and did not include potentially relevant relational or program-level variables.

Knowledge and Expectations and Trust and Communication

The model predicting Knowledge and Expectations explained 7.4% of the variance, $R^2 = .074$, adjusted $R^2 = .041$. The model predicting Trust and Communication explained 6.8% of the variance, $R^2 = .068$, adjusted $R^2 = .035$. Neither model produced statistically significant individual demographic coefficients or reference-group contrasts at $p < .05$. These results were consistent with the limited demographic differentiation identified in the nonparametric comparisons.

Home-Based Engagement

The overall OLS model predicting Home-Based Engagement was statistically significant, $F(17, 478) = 3.58$, $p < .001$, $R^2 = .113$, adjusted $R^2 = .081$. The regression coefficients, standard errors, confidence intervals, and coefficient-level p values were estimated using HC3 robust inference.

Compared with respondents with one child, respondents with two children reported lower Home-Based Engagement, $b = -0.160$, robust SE = 0.042, 95% CI [-0.243, -0.077], $p < .001$. Lower scores were also observed among respondents with three children, $b = -0.251$, robust SE = 0.049, 95% CI [-0.347, -0.155], $p < .001$, and four or more children, $b = -0.218$, robust SE = 0.053, 95% CI [-0.322, -0.114], $p < .001$.

Rural residence was associated with lower Home-Based Engagement relative to urban residence, $b = -0.092$, robust SE = 0.038, 95% CI [-0.167, -0.017], $p = .016$.

After controlling for the remaining demographic variables, none of the educational-level contrasts with the primary school reference group was statistically

significant for Home-Based Engagement. Relative to primary school education, the coefficients were $b = 0.021$, $p = .721$, for middle school education; $b = 0.047$, $p = .394$, for high school education; and $b = 0.073$, $p = .181$, for bachelor's degree and above. Thus, the educational differences observed in the univariate analysis were not supported by the adjusted reference-group comparisons.

Sensitivity Analysis Restricted to Mothers

Because the sample consisted predominantly of mothers, the multivariable regression analyses were repeated using only the 451 maternal respondents. The mother-only analyses produced substantively similar results. As in the primary regression analyses, coefficient-level standard errors and p values were based on the HC3 estimator.

Relative to mothers with one child, Home-Based Engagement remained lower among mothers with two children, $b = -0.157$, $p < .001$; three children, $b = -0.246$, $p < .001$; and four or more children, $b = -0.211$, $p < .001$. Rural residence also remained negatively associated with Home-Based Engagement, $b = -0.089$, $p = .021$.

Statistically significant educational-level contrasts remained evident for Total Parent Engagement, whereas none of the educational-level contrasts was statistically significant for Home-Based Engagement in the adjusted mother-only model. These sensitivity results indicated that the principal regression patterns were not driven by the small group of fathers. Nevertheless, the strong maternal predominance of the sample limits the generalizability of the findings to fathers.

Overall, demographic differences were limited in the Knowledge and Expectations and Trust and Communication dimensions. The clearest differences involved Home-Based Engagement, particularly in relation to family size and residential context. Respondents with one child and those living in urban areas reported higher Home-Based Engagement after the remaining demographic characteristics were controlled. However, Home-Based Engagement demonstrated low internal consistency, and its regression model explained only a limited proportion of the variance. These associations should therefore be regarded as preliminary and require replication using a psychometrically stronger measure and more representative samples.

Discussion

Table 2*Descriptive Statistics and Distributional Characteristics of Parent Engagement Scores*

Measure	N	M	SD	Median	Min	Max	Skewness	Kurtosis	Shapiro–Wilk W	p
Total Parent Engagement	496	3.73	0.21	3.80	2.92	4.00	-1.14	1.21	0.911	< .001
Knowledge and Expectations	496	3.74	0.25	3.78	2.56	4.00	-1.21	2.02	0.908	< .001
Trust and Communication	496	3.77	0.28	3.91	2.46	4.00	-1.47	2.40	0.800	< .001
Home-Based Engagement	496	3.63	0.29	3.60	2.40	4.00	-0.41	-0.48	0.902	< .001

Note. Scores were calculated as item means and ranged from 1 to 4, with higher scores indicating greater parent engagement. M = mean; SD = standard deviation; Min = minimum; Max = maximum.

Table 3*Internal Consistency and Item-Level Evidence for the Parent Engagement Scores*

Scale or subdimension	Items	Raw α	ω	Corrected item–total correlations	Mean Pearson inter-item correlation
Total Parent Engagement	25	.78	.80	.24–.56	.15
Knowledge and Expectations	9	.61	.64	.25–.48	.15
Trust and Communication	11	.77	.79	.32–.59	.24
Home-Based Engagement	5	.48	.52	.16–.38	.17

Note. Cronbach's alpha coefficients were calculated from the raw item covariance matrix. McDonald's omega coefficients were estimated using Pearson item correlations. Corrected item–total correlations and mean inter-item correlations were also based on Pearson correlations. Because raw alpha incorporates differences in item variances, it does not necessarily correspond exactly to the standardized alpha implied by the mean inter-item correlation. The low reliability of Home-Based Engagement indicates substantial measurement error; findings involving this subdimension should therefore be regarded as preliminary.

Table 4*FDR-Corrected Demographic Differences in Parent Engagement*

Outcome	Demographic variable	Test statistic	p	FDR-adjusted q	Effect size
Total Parent Engagement	Educational level	H(3) = 14.82	.002	.021	$\epsilon^2 = .030$
Home-Based Engagement	Number of children	H(3) = 37.21	< .001	< .001	$\epsilon^2 = .075$
Home-Based Engagement	Educational level	H(3) = 17.96	< .001	.002	$\epsilon^2 = .036$
Home-Based Engagement	Place of residence	U = 12,554.50	.004	.013	r = .12

Note. Only demographic group comparisons that remained statistically significant after the Benjamini–Hochberg false discovery rate correction are presented. The correction was applied separately to the 10 demographic comparisons conducted for each engagement outcome. FDR-adjusted demographic group-comparison results are reported as q values. Dunn pairwise comparisons were separately adjusted using the Benjamini–Hochberg procedure within each set of post hoc comparisons. For Kruskal–Wallis tests, ordinal epsilon-squared was calculated as $\epsilon^2 = H/(N - 1)$. Effect size r is reported for the Mann–Whitney U test. FDR = false discovery rate.

Table 5*Multiple Regression Models Predicting Parent Engagement*

Outcome and predictor	B	Robust SE	95% CI	p
Total Parent Engagement				
Middle school (ref. primary school)	-0.146	0.045	[-0.235, -0.057]	.001
High school (ref. primary school)	-0.166	0.044	[-0.253, -0.079]	< .001
Bachelor's degree and above (ref. primary school)	-0.153	0.043	[-0.238, -0.068]	< .001
Home-Based Engagement				
Two children (ref. one child)	-0.160	0.042	[-0.243, -0.077]	< .001
Three children (ref. one child)	-0.251	0.049	[-0.347, -0.155]	< .001
Four or more children (ref. one child)	-0.218	0.053	[-0.322, -0.114]	< .001
Rural residence (ref. urban residence)	-0.092	0.038	[-0.167, -0.017]	.016
Middle school (ref. primary school)	0.021	0.059	[-0.095, 0.137]	.721
High school (ref. primary school)	0.047	0.055	[-0.061, 0.155]	.394
Bachelor's degree and above (ref. primary school)	0.073	0.055	[-0.034, 0.180]	.181

Note. Unstandardized regression coefficients are reported. SE = standard error; CI = confidence interval; ref. = reference category. The reported model-level F tests are conventional OLS F tests. Coefficient-level standard errors, confidence intervals, and p values are based on the HC3 heteroskedasticity-consistent covariance estimator. Only statistically significant coefficients are presented for Total Parent Engagement. Educational-level coefficients are additionally presented for Home-Based Engagement to show that the univariate educational differences were not supported by the adjusted reference-group comparisons.

The present study aimed to examine parental engagement in early childhood education within the Turkish context by conceptualizing engagement as a multidimensional construct and investigating its associations with demographic characteristics. Overall, the findings indicate that parental engagement levels were generally high among parents of preschool-aged children; however, this high level of engagement was not uniform across dimensions. While parents reported consistently high scores in Knowledge and Expectations and Trust and Communication, more pronounced variability emerged in Home-Based Engagement. Demographic differences were largely confined to this dimension, suggesting that parental engagement in Türkiye may be broadly endorsed at the level of beliefs, expectations, and school-related communication, yet enacted unevenly when it comes to practices embedded in the home context. This pattern aligns with conceptualizations of parent engagement as a complex and context-sensitive construct shaped by both relational and structural conditions rather than a single, homogeneous behavior (Fantuzzo et al., 2013; Weiss et al., 2010; Manzon et al., 2015).

One notable feature of the findings is the overall concentration of parental engagement scores near the upper end of the scale, indicating generally high reported engagement. This pattern may be interpreted, in part, through the normative status of parent engagement in early childhood education, where engagement is widely framed as a desirable and expected parental role. Previous research has shown that conceptual ambiguity surrounding parent engagement may lead parents to interpret engagement in broad or aspirational terms, potentially contributing to elevated self-reports (Gross et al., 2020). In the Turkish context, this tendency may be further reinforced by policy discourses that increasingly emphasize parents as active partners in children's education and highlight family-school collaboration as a central principle of early childhood programs (MoNE, 2024a; MoNE, 2024b). From a theoretical perspective, the prominence of high engagement scores is also consistent with frameworks that position program-family partnerships as a core and valued component of early childhood education systems (Halgunseth et al., 2009). At the same time, the clustering of scores near the scale ceiling suggests that reported engagement may reflect not only actual practices but also socially desirable responding shaped by institutional expectations and dominant educational narratives.

Parental engagement did not manifest uniformly across its dimensions, and the observed differentiation among subdimensions provides an important insight into how engagement is enacted in early childhood education contexts. In the present study, Knowledge

and Expectations and Trust and Communication were characterized by consistently high scores, whereas Home-Based Engagement was comparatively lower. This pattern may be understood by considering the differing visibility and structure of these engagement forms. School-related dimensions of engagement, such as communication with teachers and alignment with educational expectations, are often more explicitly defined, institutionally encouraged, and socially recognized. In contrast, home-based engagement encompasses practices that are less observable, less standardized, and more dependent on families' individual circumstances. Prior research has similarly emphasized that parent engagement comprises distinct dimensions that may not develop in parallel and may respond differently to contextual constraints (Gross et al., 2022). From a broader perspective, this distinction also mirrors the differentiation between center-oriented practices and child-focused practices, with the latter often requiring greater initiative, sustained effort, and contextual adaptation within the home environment (OECD, 2011). In line with the Social Exchange Model of Family Engagement, the enactment of home-based practices is particularly sensitive to family-level resources, suggesting that high engagement in school-related domains does not automatically translate into equally strong engagement at home (Halgunseth et al., 2009).

The clearest demographic difference in Home-Based Engagement concerned the number of children in the household. Respondents with one child reported higher scores than respondents with two, three, or four or more children. However, the multiple-child groups did not differ significantly from one another. The result should therefore be interpreted as a distinction between one-child and multiple-child households rather than as evidence of a progressive decline with each additional child. One possible explanation is that home-based practices require time, attention, and opportunities for individual interaction, which may be distributed across several children in larger households. This interpretation is compatible with the Social Exchange Model's emphasis on finite family-level resources (Halgunseth et al., 2009). Turkish research has likewise identified time constraints, caregiving responsibilities, and competing family demands as potential barriers to sustained participation (Çakmak, 2010; Orçan Kaçan et al., 2019). Nevertheless, these mechanisms were not directly measured in the present study and should be examined in future research.

In addition to family structure, residential context emerged as a meaningful factor in shaping home-based engagement. Parents residing in urban areas reported higher levels of home-based engagement compared to those living in rural settings, although the magnitude of this difference was modest. One possible explanation is that residential settings differ in access

to educational materials, informational support, and community-based learning opportunities. Recent qualitative evidence from rural Türkiye has documented spatial inequalities in families' experiences of preschool access and adequacy (Dikmen & Gökalp, 2026). However, the present study did not directly measure resource availability, service access, or community support. The observed rural-urban difference should therefore be interpreted as a contextual association rather than evidence of a specific explanatory mechanism. From a theoretical standpoint, the Social Exchange Model emphasizes the role of program-community partnerships and the broader ecological context in enabling family engagement (Halgunseth et al., 2009). When such supports are limited or less accessible, as may be the case in some rural settings, families may face additional barriers to sustaining engagement practices beyond the school context. Within Türkiye, previous research has similarly pointed to persistent challenges in translating policy-level emphasis on parent involvement into everyday practices, particularly in contexts marked by structural and resource-related constraints (Filik Uyanık, 2018). Taken together, the urban-rural differences observed in this study underscore the importance of contextual sensitivity when interpreting parental engagement and caution against assuming uniform opportunities for engagement across residential settings (OECD, 2011).

Educational level produced different patterns across the total score and Home-Based Engagement. For Total Parent Engagement, the adjusted contrasts indicated lower scores among the middle school, high school, and bachelor's-degree-and-above groups than among the primary school reference group. In contrast, the univariate Home-Based Engagement analysis showed higher scores in the bachelor's-degree-and-above group than in the middle and high school groups, but none of the educational-level contrasts remained significant in the adjusted model. These findings do not support a simple linear relationship between education and engagement. They may reflect differences among engagement dimensions, correlations between education and other demographic characteristics, or response patterns across educational groups. Because these explanations were not tested directly, the educational findings should be interpreted descriptively and examined further in future research.

The multivariable regression analyses further contribute to the interpretation of these findings by demonstrating that demographic characteristics explained only a modest proportion of variance in parental engagement outcomes. Although statistically significant adjusted educational contrasts were identified for Total Parent Engagement, and family-size and residence contrasts were identified

for Home-Based Engagement, the models explained only modest proportions of variance. The limited explanatory power indicates that demographic characteristics capture only part of the variation in parent engagement. Engagement in early childhood education is widely conceptualized as a relational and contextually embedded process shaped by interactions among families, children, educational institutions, and communities (Halgunseth et al., 2009; Weiss et al., 2010). As such, it cannot be fully accounted for by demographic indicators alone. Factors such as program quality, teacher practices, parent-teacher relationships, cultural norms, and individual beliefs are likely to play a substantial role in shaping engagement, yet these variables extend beyond the scope of demographic modeling. Previous syntheses of the literature similarly emphasize that demographic characteristics tend to exert small and indirect effects on parent engagement when considered alongside relational and institutional factors (Gross et al., 2022). Accordingly, the modest explanatory power is consistent with the view that parent engagement is shaped by demographic, relational, institutional, and cultural conditions operating together.

Finally, the present study makes a substantive contribution to the literature on early childhood education in Türkiye by addressing a longstanding gap in empirical research on parent engagement. Despite increasing policy-level emphasis on family participation in education, the concept of engagement in the Turkish context has largely been framed and examined through the lens of parent involvement, with limited attention to its multidimensional and relational nature (Kaya, 2023; Kaya, 2024). By adopting a validated, multidimensional measure of parent engagement and examining demographic patterns within a relatively large convenience sample, this study provides one of the first detailed quantitative examinations of parental engagement in early childhood education in Türkiye. In doing so, it moves beyond descriptive accounts of involvement to capture differentiated engagement practices across home and school contexts. The findings also respond to calls in the national literature for more nuanced and evidence-based examinations of parent engagement, particularly in relation to persistent structural and contextual challenges (Filik Uyanık, 2018; Serdaroğlu & Bekir, 2024). Taken together, the study offers initial empirical evidence for future research and practice by clarifying how parental engagement is distributed across dimensions and demographic contexts in Türkiye, thereby strengthening the conceptual and empirical grounding of this field.

Conclusion

This study examined parental engagement in early childhood education in Türkiye as a multidimensional

construct encompassing Knowledge and Expectations, Trust and Communication, and Home-Based Engagement. Overall, parents reported high levels of engagement, particularly in school-facing dimensions such as expectations, knowledge, and communication. However, engagement was not uniform across domains. The most meaningful variation was concentrated in Home-Based Engagement, which was comparatively lower than the other subdimensions and showed systematic differences across key demographic characteristics. In both univariate and multivariable analyses, home-based engagement tended to be lower among families with more children and among those living in rural settings, suggesting that engagement enacted in the home environment may be more sensitive to structural constraints and contextual opportunities than engagement expressed through beliefs, expectations, and school communication. Together, these findings support the view that parental engagement is broadly endorsed in principle but unevenly supported in practice, particularly in the home context. However, the Home-Based Engagement score demonstrated low internal consistency, and the corresponding associations should be treated as preliminary. The findings indicate one-child versus multiple-child and urban versus rural differences rather than establishing the mechanisms underlying these patterns.

The results suggest that early childhood education policies and program-level practices in Türkiye may benefit from shifting from general encouragement of parent engagement toward more targeted support for Home-Based Engagement, which appears to be the dimension in which the clearest demographic differences were observed in this sample. First, early childhood programs may consider integrating low-burden, high-feasibility home learning routines into their family outreach efforts. Practical tools such as short activity menus, weekly “micro-practices” (e.g., shared reading, everyday math talk, structured play prompts), and examples of learning opportunities embedded in daily routines can make home-based engagement more accessible, particularly for families experiencing time and resource constraints.

Second, the observed difference between one-child and multiple-child households underscores the need for supports that acknowledge competing demands within multi-child households. Programs may consider differentiated guidance for parents with multiple children, such as strategies for managing shared learning activities across siblings, rotating one-to-one interaction time, or using group-based home activities that simultaneously support multiple children. Designing materials that explicitly address “how to do this with more than one child” can increase feasibility and reduce the implicit assumption that engagement requires intensive one-to-one time.

Third, the rural–urban disparity in home-based engagement highlights the importance of improving access to family support mechanisms in rural contexts. This does not necessarily require resource-intensive interventions. Scalable options such as mobile-friendly guidance, short video demonstrations, parent messaging groups facilitated by teachers, and community-based partnerships that expand access to books and learning materials may help reduce contextual barriers. Importantly, these strategies should be implemented in culturally responsive ways and framed as supportive rather than evaluative, given that home-based engagement is closely tied to families’ everyday realities.

Finally, the limited demographic differentiation in Knowledge and Expectations and Trust and Communication suggests that these scores were relatively homogeneous in the present sample. However, the strong ceiling effects mean that this homogeneity may partly reflect restricted score variability. Existing school–family communication channels may nevertheless provide a practical route for offering optional guidance on home-based learning practices.

Limitations

Several limitations should be considered. The most important concerns the measurement of Home-Based Engagement, which constituted the main focus of several findings. This five-item subdimension demonstrated low internal consistency, with $\alpha = .48$ and $\omega = .52$. Corrected item–total correlations ranged from .16 to .38, and its standardized factor loadings ranged from .31 to .59. Although the ordinal CFA supported the overall three-factor structure, acceptable global model fit does not eliminate the subdimension-level reliability problem. The clustering of responses in the highest category may have restricted score variability and weakened inter-item relationships. Consequently, Home-Based Engagement scores contain substantial measurement error, and the demographic associations involving this outcome may be attenuated, unstable, or sample-specific. These findings should therefore be regarded as preliminary rather than definitive. A second limitation concerns the strong maternal predominance of the sample. Mothers constituted 90.9% of participants, meaning that the findings primarily represent maternal reports and cannot be generalized confidently to fathers. Although the mother-only sensitivity analysis produced substantively similar regression patterns, it does not compensate for the limited representation of fathers. Third, convenience sampling with snowball features and online recruitment limits the representativeness of the sample. Parents with internet access and connections to the networks through which the questionnaire was distributed may differ from other parents in Türkiye.

The urban predominance of the sample further limits broader generalization. Fourth, the study relied on self-reported engagement. Because engagement is socially valued, responses may have been influenced by social desirability. The pronounced ceiling effects also reduced variability and may have limited the detection of demographic differences. The moderate internal consistency of Knowledge and Expectations ($\alpha = .61$; $\omega = .64$) should likewise be considered when interpreting findings for that dimension. Fifth, the cross-sectional design supports associations but not causal or directional conclusions. Explanations involving time constraints, resource availability, self-evaluation, or access to educational opportunities were not directly tested. Finally, the regression models explained only modest proportions of variance and did not include potentially important relational and program-level factors, such as teacher outreach, school climate, parent-teacher trust, family routines, or access to learning resources.

Future Research Avenues

Future research can extend this work in several ways. First, longitudinal designs are needed to examine how parent engagement evolves over time and how demographic and contextual factors shape trajectories of engagement across children's early education experiences. Such designs would also allow stronger inferences regarding directionality between family characteristics and engagement practices. Second, mixed-methods studies could deepen understanding of why home-based engagement varies across family size and residential context. Qualitative interviews or focus groups with parents and teachers may clarify specific barriers, culturally grounded meanings of "engagement," and the types of supports families perceive as most feasible and respectful. Third, future research should incorporate program- and relationship-level predictors alongside demographics, such as teacher outreach practices, perceived school climate, parent-teacher trust, and access to learning resources. Including these variables may improve explanatory power and better reflect the multilevel nature of parent engagement. Fourth, measurement-focused research should re-examine the Home-Based Engagement items in more diverse Turkish samples. Future studies should evaluate whether the weak reliability reflects item heterogeneity, restricted variance, contextual interpretation, or insufficient coverage of home-based practices. Item revision or expansion, test-retest reliability, and replication of the factor structure should be considered before strong substantive conclusions are drawn from this subdimension. Fifth, future samples should include more balanced numbers of mothers and fathers and greater representation of rural families. Such designs would permit more credible examination of whether engagement patterns differ according to parent gender and residential context.

Finally, once the Home-Based Engagement findings have been replicated using psychometrically stronger measures, intervention studies may evaluate low-cost strategies designed to support home-based practices in multiple-child households and rural settings.

Declarations

Acknowledgments

The authors would like to thank all parents who voluntarily participated in the study.

Funding

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Ethics Statements

Ethical approval for this study was obtained from the Afyon Kocatepe University Social and Human Sciences Scientific Research and Publication Ethics Committee (Approval No: 2024/70). The study was conducted in accordance with the ethical standards for research involving human participants.

Conflict of Interest

The authors declare that they have no conflict of interest.

Informed Consent

Informed consent was obtained electronically from all participants prior to participation. Participation was voluntary, and responses were collected anonymously.

Data availability

The datasets generated and/or analyzed during the current study are available from the corresponding author on reasonable request.

AI Disclosure Statement

The authors used ChatGPT (OpenAI, GPT-5.5) solely to assist with English language editing, grammar correction, and improving the clarity of the manuscript. The AI tool was not used to generate the study design, analyses, results, interpretations, or scientific conclusions. All AI-assisted edits were carefully reviewed, verified, and, where necessary, revised by the authors, who take full responsibility for the final content of the manuscript.

References

- Atış Akyol, N., Doğan, S. A., & Durmuşoğlu, M. (2021). Okul öncesi eğitimde öğretmenin iletişim becerilerinin önemi: Aile katılımı, iş birliği ve öğretmen-çocuk ilişkisi [The importance of teacher's communication skills in preschool education: family involvement, cooperation and teacher-child relationship]. *Bolu Abant İzzet Baysal University Journal of Faculty of Education*, 21(3), 887-899. <https://doi.org/10.17240/aibuefd.2021.21.64908-900473>
- Barnett, M. A., Paschall, K. W., Mastergeorge, A. M., Cutshaw, C. A., & Warren, S. M. (2020). Influences of parent engagement in early childhood education centers and the home on kindergarten school readiness. *Early Childhood Research Quarterly*, 53, 260-273. <https://doi.org/10.1016/j.jecresq.2020.05.005>
- Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society: Series B (Methodological)*, 57(1), 289-300. <https://doi.org/10.1111/j.2517-6161.1995.tb02031.x>
- Bettencourt, A. F., Gross, D., Bower, K., Francis, L., Taylor, K., Singleton, D. L., & Han, H. R. (2023). Identifying meaningful indicators of parent engagement in early learning for low-income, urban families. *Urban Education*, 58(10), 2308-2345. <https://doi.org/10.1177/0042085920968619>
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Lawrence Erlbaum.
- Cohen, L., Manion, L., & Morrison, K. (2018). *Research methods in education* (8th ed.). Routledge.
- Creswell, J. W., & Creswell, J. D. (2018). *Research design: Qualitative, quantitative, and mixed methods approaches* (5th ed.). Sage.
- Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, 16(3), 297-334.
- Cutshaw, C. A., Mastergeorge, A. M., Barnett, M. A., & Paschall, K. W. (2022). Parent engagement in early care and education settings: Relationship with engagement practices and child, parent, and centre characteristics. *Early Child Development and Care*, 192(3), 442-457. <https://doi.org/10.1080/03004430.2020.1764947>
- Çakmak, Ö. Ç. (2010). Okul öncesi eğitim kurumlarında aile katılımı [Parent involvement in pre-school]. *Bolu Abant İzzet Baysal University Journal of Social Sciences*, 10(1), 1-18. <https://doi.org/10.11616/AbantSbe.256>
- Demircan, H. Ö. (2018). Okul öncesi eğitimde aile katılımı sistemini genişletmek: Aile "bağlılığı, ortaklığı, katılımı ve eğitimi" [Expanding the Parent Involvement Framework in Early Childhood Education: Parent "Engagement, Partnership, Involvement and Education"]. *İlköğretim Online*, 17(4), 1-19. <https://doi.org/10.17051/ilkonline.2019.507008>
- Dikmen, M., & Gökalp, A. C. (2026). Rethinking preschool education in rural Turkey: Parents' lived experiences of access, quality, and inequality. *Early Childhood Education Journal*. Advance online publication. <https://doi.org/10.1007/s10643-026-02177-w>
- Epstein, J. L. (1995). School/family/community partnerships: Caring for the children we share. *Phi Delta Kappan*, 76(9), 701-712.
- Erdem, A. R. (2005). İlköğretimimizin gelişimi ve bugün gelinen nokta [The development of our primary education and where we stand today]. *Bilim, Eğitim ve Düşünce Dergisi*, 5(2), 1-15.
- Fantuzzo, J., Gadsden, V., Li, F., Sproul, F., McDermott, P., Hightower, D., & Minney, A. (2013). Multiple dimensions of family engagement in early childhood education: Evidence for a short form of the Family Involvement Questionnaire. *Early Childhood Research Quarterly*, 28(4), 734-742. <https://doi.org/10.1016/j.jecresq.2013.07.001>
- Ferlazzo, L. (2011). Involvement or engagement. *Educational Leadership*, 68(8), 10-16.
- Field, A. P. (2013). *Discovering statistics using IBM SPSS Statistics* (4th ed.). Sage.
- Filik Uyanık, R. (2018). *Motivators of mothers' engagement in their preschoolers' education regarding the influence of media* (Master's thesis). Middle East Technical University, Ankara, Turkey. YÖK Thesis Center. <https://tez.yok.gov.tr/UlusalTezMerkezi/tezSorguSonucYeni.jsp>
- Forry, N., Bromer, J., Chrisler, A., Rothenberg, L., Simkin, S., & Daneri, P. (2012). *Family-provider relationship quality: Review of conceptual and empirical literature of family-provider relationships*. PsycEXTRA Dataset. <https://doi.org/10.1037/e537912013-001>

- Fraenkel, J. R., Wallen, N. E., & Hyun, H. H. (2019). *How to design and evaluate research in education* (10th ed.). McGraw-Hill.
- Gitonga, C. M., Karuku, S., Asweto, C. O., Rugano, P., Saoko, V. O., Mulonzi, B. M., Nzomo, C. M., Mwenda, E., & Ngichio, D. O. (2026). *Determinants of parental engagement in early childhood education in Kenya*. *Discover Education*, 5, 56. <https://doi.org/10.1007/s44217-025-01059-4>
- Goodall, J., & Montgomery, C. (2014). Parental involvement to parental engagement: A continuum. *Educational Review*, 66(4), 399–410. <https://doi.org/10.1080/00131911.2013.781576>
- Gross, D., Bettencourt, A. F., Finch, W. H., Plesko, C., Paulson, R., & Singleton, D. L. (2022). Developing an equitable measure of parent engagement in early childhood education for urban schools. *Children and Youth Services Review*, 141, 106613. <https://doi.org/10.1016/j.childyouth.2022.106613>
- Gross, D., Bettencourt, A. F., Taylor, K., Francis, L., Bower, K., & Singleton, D. L. (2020). What is parent engagement in early learning? Depends who you ask. *Journal of Child and Family Studies*, 29(3), 747–760. <https://doi.org/10.1007/s10826-019-01680-6>
- Halgunseth, L. C., Peterson, A., Stark, D. R., & Moodie, S. (2009). *Family engagement, diverse families, and early childhood education programs: An integrated review of the literature*. National Association for the Education of Young Children; The Pew Charitable Trust.
- Halgunseth, L., Jia, G., & Barbarin, O. (2013). Family engagement in early childhood programs: Serving families of dual language learners. In L. Espinosa (Ed.), *California's best practices for young dual language learners* (pp. 119–171). CDE Publications.
- Harvard Family Research Project. (2014). *Redefining family engagement for student success*. https://www.studenthealthandlearning.org/wp-content/uploads/2016/05/Redefining-Engagement_May2014.pdf
- Head Start. (2014). *Family engagement as parent involvement 2.0: Understanding the difference in terms & concepts*. <https://eclkc.ohs.acf.hhs.gov/sites/default/files/pdf/fam-engage-parent-inv.pdf>
- Kağıtçıbaşı, C., Sunar, D., Bekman, S., Baydar, N., & Cemalcılar, Z. (2009). Continuing effects of early enrichment in adult life: The Turkish Early Enrichment Project 22 years later. *Journal of Applied Developmental Psychology*, 30(6), 764–779. <https://doi.org/10.1016/j.appdev.2009.05.003>
- Karasar, N. (2022). *Bilimsel araştırma yöntemi* [Scientific research method] (37. baskı). Nobel Publishing.
- Kaya, Ü. Ü. (2023). *Parental satisfaction and engagement in early childhood education: The mediating role of home literacy environment* (Doctoral dissertation). Middle East Technical University.
- Kaya, Ü. Ü. (2024). Türkiye'deki okul öncesi eğitim programlarında aile katılımının evrimi [The evolution of family participation in early childhood education programs in Turkey]. In I. Erpay & E. R. O. Ağae (Eds.), *Proceedings of the 2nd Bilisel International Gordion Scientific Researches Congress* (pp. 397–415). Astana Publications.
- Kaya, Ü. Ü., Demircan, H. Ö., & Altun, D. (2024). Adaptation and validation of the Parent Engagement in Early Childhood Education Survey for the Turkish context. *HAYEF: Journal of Education*, 21(3), 244–251. <https://doi.org/10.5152/hayef.2024.23096>
- Kruskal, W. H., & Wallis, W. A. (1952). Use of ranks in one-criterion variance analysis. *Journal of the American Statistical Association*, 47(260), 583–621. <https://doi.org/10.1080/01621459.1952.10483441>
- Long, J. S., & Ervin, L. H. (2000). Using heteroscedasticity-consistent standard errors in the linear regression model. *The American Statistician*, 54(3), 217–224. <https://doi.org/10.1080/00031305.2000.10474549>
- Mann, H. B., & Whitney, D. R. (1947). On a test of whether one of two random variables is stochastically larger than the other. *The Annals of Mathematical Statistics*, 18(1), 50–60. <https://doi.org/10.1214/aoms/1177730491>
- Manzon, M., Miller, R., Hong, H., & Khong, L. (2015). Parent Engagement in Education (NIE Working Paper Series No. 7). Singapore: National Institute of Education.
- Ministry of National Education. (1989). *Okul öncesi öğretmen kılavuz kitabı* [Pre-School Teacher Guide Book]. Milli Eğitim.

- Ministry of National Education. (1994). *Preschool Education Curriculum (37–60 Months)*. Milli Eğitim.
- Ministry of National Education. (2002). *Preschool Education Curriculum for 36–72 Months Old Children*. Ya-Pa.
- Ministry of National Education. (2006). *Early Childhood Education Program*.
- Ministry of National Education. (2013). *Early Childhood Education Program*. General Directorate of Basic Education.
- Ministry of National Education. (2024a). *Early Childhood Education Program* [<https://tegm.meb.gov.tr/dosya/okuloncesi/guncellenenokuloncesiegitimprogrami.pdf>]
- Ministry of National Education. (2024b). *Early Childhood Education Program Türkiye Century Education Model* [https://tegm.meb.gov.tr/meb_iys_dosyalar/2024_09/20104013_2024programokuloncesionayli.pdf]
- Oakes, J., & Lipton, M. (2007). *Teaching to change the world*. McGraw-Hill.
- OECD. (2011). Engaging families and communities. In *Starting strong III: A quality toolbox for early childhood education and care*. <https://doi.org/10.1787/9789264123564-7-en>
- Orçan Kaçan, M., Kimzan, İ., Güler Yıldız, T., & Çağdaş, A. (2019). Öğretmen ve ebeveynlerin aile katılımını etkileyen etmenlere yönelik görüşlerinin incelenmesi [Investigation of teacher and parent views on factors affecting family participation]. *Journal of Education, Theory and Practical Research*, 5(3), 370-381.
- Serdaroğlu, H. U., & Bekir, H. (2024). Okul öncesi eğitim programında güncellemeler [Updates in preschool education program]. *Journal of Individual & Society*, 14(1), 1–29. <https://doi.org/10.20493/birtop.1489551>
- Shapiro, S. S., & Wilk, M. B. (1965). An analysis of variance test for normality (complete samples). *Biometrika*, 52(3–4), 591–611. <https://doi.org/10.1093/biomet/52.3-4.591>
- Tekin, A. K. (2011). Parent involvement revisited: Background, theories, and models. *International Journal of Applied Educational Studies*, 11, 1–13.
- Tomczak, M., & Tomczak, E. (2014). The need to report effect size estimates revisited: An overview of some recommended measures of effect size. *Trends in Sport Sciences*, 21(1), 19–25.
- United States Department of Education. (2015). *Every Student Succeeds Act (ESSA)*. United States Department of Education.
- Weiss, H. B., Kreider, H., Lopez, M. E., & Chatman-Nelson, C. (2010). *Preparing educators to engage families: Case studies using an ecological systems framework* (2nd ed.). SAGE.
- Wondim, M. G. (2025). Parental engagement in early childhood education: The case of Debre Tabor schools, Ethiopia. *European Journal of Education*, 60(1), e12817. <https://doi.org/10.1111/ejed.12817>
- Varol, V., Kılıç, M., Yılmaz Çelik, H., Güneş, M., & Varol, M. (2023). Okul öncesi dönemde anne ve babaların aile katılım çalışmalarına katılım düzeylerinin incelenmesi [Examining mothers' and fathers' participation levels in family involvement activities during the preschool period]. *International Journal of Humanities and Education*, 9(19), 157–177. <https://doi.org/10.59304/ijhe.1219001>

The Development and Validation of the Metacognitive Awareness Scale for Middle School Students*

Ayşe Gönültaş Yıldırım^a, Ayşegül İşlekeller Bozca^b

Received : 18 August 2026
Revised : 6 September 2026
Accepted : 30 September 2026
DOI : 10.26822/iejee.2026.466

^a **Corresponding Author:** Ayşe Gönültaş Yıldırım, Social Services Specialist, Ministry of Interior of the Republic of Türkiye
E-mail: gonultasayse@gmail.com
ORCID: <https://orcid.org/0000-0003-4354-1723>

^b Ayşegül İşlekeller Bozca, Istanbul University-Cerrahpaşa, Department of Special Education
E-mail: islekellerbozca@iuc.edu.tr
ORCID: <https://orcid.org/0000-0003-2912-5379>

* This article was derived from the doctoral dissertation of the first author, Ayşe Gönültaş Yıldırım, under the supervision of Ayşegül İşlekeller Bozca.

Abstract

This study aimed to develop a psychometrically sound instrument to assess the metacognitive awareness of middle school students. A 5-point Likert-type instrument was developed to assess middle school students' metacognitive knowledge and regulatory processes. The study was conducted with a total of 1009 students. Exploratory factor analysis (EFA) was conducted with data from 475 students. Preliminary diagnostics supported factorability (KMO = .952; Bartlett's $\chi^2 = 8377.999$, $p < .01$). Based on statistical and theoretical criteria, 17 items were removed, yielding a unidimensional structure consisting of 29 items that explained 38.392% of the total variance. Factor loadings ranged between .469 and .644. Confirmatory factor analysis (CFA) was subsequently performed with an independent sample of 534 students. Model fit indices ($\chi^2/df = 2.770$, CFI = .904, TLI = .899, IFI = .904, NFI = .870, RFI = .860, RMSEA = .048) indicated an acceptable fit for the proposed one-factor model. Internal consistency was high (Cronbach's $\alpha = .942$), and item-total correlations ranged from .467 to .651. Test-retest reliability, assessed with 126 students from a public middle school over a three-week interval, demonstrated excellent temporal stability (ICC(3,1) = .997, 95% CI [.996, .998]). Overall, the scale demonstrated strong validity and reliability and provides a developmentally appropriate instrument for assessing metacognitive awareness in middle school students.

Keywords:

Metacognition, Middle School Students, Scale Development, Validity and Reliability, Temporal Stability

Introduction

Metacognition refers to individuals' awareness and regulation of their own cognitive processes (Flavell, 1979; Schraw & Dennison, 1994; Veenman et al., 2006). It comprises two fundamental components: metacognitive knowledge and metacognitive regulation (Baker & Brown, 1980; Pintrich, 2002; Schraw & Dennison, 1994; Schraw & Moshman, 1995; Vrugt & Oort, 2008). Although metacognition has been conceptualized in different ways, most contemporary frameworks distinguish between knowledge of cognition and regulation of cognition (Schraw & Dennison, 1994;



www.iejee.com
ISSN: 1307-9298

2026 Published by KURA Education & Publishing.
This is an open access article under the CC BY-NC-ND license. (<https://creativecommons.org/licenses/by/4.0/>)

Veenman et al., 2006). Based on these theoretical perspectives, the conceptual framework of metacognition is presented in Figure 1.

Figure 1.

Conceptual framework of metacognition synthesized from previous theoretical models.



Metacognitive knowledge refers to individuals' understanding of their own cognitive processes and includes declarative, procedural, and conditional knowledge (Cross & Paris, 1988; Schraw & Moshman, 1995). Metacognitive regulation refers to the processes through which learners plan, monitor, control, and evaluate their cognitive activities during learning (Jacobs & Paris, 1987; Schraw & Moshman, 1995; Veenman et al., 2006). Together, these complementary components form the theoretical basis of metacognitive awareness and provide the conceptual framework underlying the development of the present scale.

Students with well-developed metacognitive skills regulate their learning more effectively by planning, monitoring, and evaluating their cognitive processes (Kuhn, 2000). These competencies are particularly important during middle school, when students experience rapid cognitive development and establish increasingly independent learning habits (Karakelle & Saraç, 2010; Özsoy, 2011). Consistent with this developmental perspective, previous research has shown that higher levels of metacognitive awareness are associated with improved academic achievement, motivation, problem-solving, and self-regulated learning (Coutinho & Neuman, 2008; Desoete, 2008; Howard et al., 2000; Jahangard et al., 2016; Vrugt & Oort, 2008). Furthermore, intervention studies have demonstrated that metacognitive skills can be enhanced through appropriately designed instructional approaches, highlighting the importance of educational practices that explicitly foster metacognitive awareness (Çakıroğlu, 2007; Güner & Erbay, 2021; Owo & Ikwut, 2015; Özsoy, 2011; Young & Fry, 2008).

As students transition from concrete to formal operational thinking during middle school, assessing their metacognitive awareness becomes increasingly important for supporting independent learning and self-regulation. Reliable assessment instruments are therefore needed to help teachers identify students' metacognitive strengths and weaknesses and to design appropriate instructional interventions. A

review of the literature on metacognition reveals that several measurement instruments have been developed internationally for children (Annevirta et al., 2007; Jacobs & Paris, 1987; Lecce et al., 2010; Mokhtari & Reichard, 2002; Pereira-Laird & Deane, 1997; Schraw & Dennison, 1994; Sperling et al., 2002). However, in Türkiye, available instruments assessing metacognition remain limited (Abacı et al., 2006; Çetinkaya & Erkin, 2002; Karakelle & Saraç, 2007; Öztürk, 2012).

Among the earliest Turkish instruments, Çetinkaya and Erkin (2002) developed a scale examining metacognition in relation to reading comprehension and academic achievement. However, the instrument focused primarily on reading strategies rather than metacognition as a broader construct. Similarly, Karakelle and Saraç (2007) developed the Metacognitive Awareness Scale for Children for students aged 10–14. Although this scale represented an important contribution, it primarily assessed declarative awareness and provided limited coverage of regulatory processes such as planning, monitoring, and evaluation.

In addition to these national instruments, one of the most widely used international tools is the Metacognitive Awareness Inventory (MAI) developed by Schraw and Dennison (1994). However, the MAI was originally designed for adults and university students, and its level of abstraction and linguistic complexity may limit its suitability for middle school students. Similarly, the Junior Metacognitive Awareness Inventory (Jr. MAI), although developed for younger populations, has been reported to focus more on declarative knowledge and does not sufficiently capture the dynamic and process-oriented nature of regulatory processes such as planning, monitoring, and evaluation.

Furthermore, many existing instruments, both internationally and in the Turkish context, tend to assess metacognition either within specific domains (e.g., reading) or within limited dimensions, and may not adequately reflect the integrated and process-based nature of metacognitive functioning in contemporary learning environments. Given these limitations, there is a clear need to develop a comprehensive, up-to-date, and developmentally appropriate metacognitive awareness scale for middle school students. Such an instrument should encompass both metacognitive knowledge and control processes while accounting for students' cognitive, linguistic, and cultural characteristics.

Collectively, previous instruments either target different age groups, focus on domain-specific metacognitive skills, or fail to integrate metacognitive knowledge and regulatory processes within a single developmentally appropriate framework. Consequently, a psychometrically robust instrument specifically

designed for middle school students remains unavailable. No existing instrument simultaneously captures both metacognitive knowledge and regulatory processes in a developmentally appropriate, process-oriented, and culturally relevant manner for middle school students within the Turkish educational context. Developing this type of scale is therefore a critical priority for supporting effective learning and instructional design at the middle school level. The present study aims to address this gap by developing a metacognitive awareness scale that integrates both knowledge and control processes, reflects the dynamic and process-based nature of learning, is linguistically and cognitively appropriate for middle school students, and aligns with contemporary educational frameworks emphasizing self-regulated learning.

Development of the Scale and Significance of the Study

Middle school students are in a developmental period marked by the transition from concrete to abstract thinking. At this stage, their skills in self-expression and self-evaluation may still be limited. Therefore, the wording, conceptual level, and questioning style of metacognition scales developed for high school or university students may be insufficient in terms of clarity and validity for middle school learners. In addition, when considering the cultural context, some existing scales require reconstruction to reflect the social, cognitive, and linguistic realities of middle school students in Turkey.

This situation highlights the need for a new, valid, and reliable measurement tool that accurately reflects the fundamental components of metacognition and is suitable for the age, developmental level, and cultural context of middle school students. A scale developed for this purpose would not only help determine students' levels of metacognitive awareness but also assist teachers in guiding learning processes more effectively. Moreover, in line with the constructivist learning principles of the Türkiye Yüzyılı Maarif Model, such a tool would contribute to the literature by providing a contemporary instrument that comprehensively measures students' abilities to construct, plan, monitor, and evaluate knowledge.

The purpose of this study is to develop a valid and reliable metacognitive awareness scale for middle school students that is tailored to their cognitive and linguistic characteristics, encompassing the primary metacognitive processes: cognitive awareness, planning, monitoring, regulation, and evaluation. This study aims to address the lack of appropriate scales in the national literature and to make both theoretical and practical contributions by providing a tool that enables students to manage their own learning processes more intentionally. In addition to addressing an important measurement gap, the proposed scale

provides researchers and educators with a practical instrument for evaluating students' metacognitive awareness within contemporary educational settings.

Method

Research Design and Sample

This study employed a general survey design aimed at developing a metacognitive awareness scale. In research based on survey models, a past or existing phenomenon, event, or situation is described as it naturally occurs without any external manipulation (Büyüköztürk et al., 2009). The survey model is typically used to identify individuals' opinions or characteristics such as interests, abilities, and attitudes, and it is generally conducted with larger samples than other research designs (Büyüköztürk et al., 2009).

A purposive sampling method was used in the study. Purposive sampling is preferred in cases where a detailed investigation is necessary and allows for selecting participants who are both information-rich and capable of representing the target population, based on available resources (Büyüköztürk et al., 2009). Since the research sample was selected according to specific criteria such as district, school type, and grade level, the purposive sampling method was considered appropriate (Büyüköztürk et al., 2009). In scale development studies, determining the sample size is based on theoretical principles. It is generally recommended that the sample size be at least five times the number of items on the scale (Conway & Huffcutt, 2003). While 200 participants are considered sufficient to obtain reliable factors in factor analysis (Tabachnick & Fidell, 2007), 300 participants are regarded as good, 500 as very good, and 1000 as excellent (Kline, 1998).

The study group consisted of students enrolled in four different middle schools located in the Esenyurt, Avcılar, and Beylikdüzü districts of Istanbul during the first semester of the 2025–2026 academic year. Before data collection, ethical approval was obtained from the Ethics Committee of Istanbul University, along with permissions from the Ministry of National Education (MEB), the participating institutions, teachers, and parents. The data were collected face-to-face through paper-based questionnaires. A total of 1,067 students participated in the study; however, data from 58 students were excluded from the analysis due to various reasons. These excluded cases included 19 fifth-grade, 15 sixth-grade, 11 seventh-grade, and 13 eighth-grade students. The primary reasons for exclusion were incomplete responses, multiple selections for a single item, and answering patterns that indicated low validity (e.g., marking only "strongly agree" for all items). The analyses were conducted using valid data from 1,009 students. Of these participants, 494 (49%) were female and 515 (51%) were male. The distribution by grade level was as follows: fifth grade – 297 (29%),

sixth grade – 263 (26%), seventh grade – 231 (23%), and eighth grade – 218 (22%). In terms of school type, 577 (57%) students attended public middle schools, 229 (23%) attended imam hatip middle schools, and 203 (20%) attended private middle schools. The developed metacognitive awareness scale was administered across these diverse populations, and the validity and reliability of the scale were examined accordingly.

Table 1.

Descriptive Characteristics of Middle School Students

Descriptive Variable	Category	f	%
Gender	Female	494	49
	Male	515	51
Grade Level	5th Grade	297	29
	6th Grade	263	26
	7th Grade	231	23
	8th Grade	218	22
School Type	Imam Hatip Middle School	229	23
	Public Middle School	577	57
	Private Middle School	203	20
Total		1009	100

Additionally, data were recollected for evaluation of test–retest reliability. However, since it was not feasible to obtain data from all participants again, this stage involved only 126 students from a public middle school. Among these participants, 48.4% ($n = 61$) were female and 51.6% ($n = 65$) were male. Regarding grade distribution, 26.2% ($n = 33$) were fifth-grade students, 27% ($n = 34$) were sixth-grade students, 22.2% ($n = 28$) were seventh-grade students, and 24.6% ($n = 31$) were eighth-grade students. All students in this group attended public middle schools, with a total of 126 participants.

Development of the Metacognitive Awareness Scale for Middle School Students

Metacognition is defined as an individual's ability to recognize, monitor, evaluate, and, when necessary, regulate their own cognitive processes (Flavell, 1979). This concept indicates that learning is not limited to the acquisition of knowledge but also involves reflecting on how learning occurs. Metacognitive skills encompass the dimensions of cognitive awareness, planning, monitoring, evaluation, and regulation, enhancing students' ability to manage their learning processes more effectively (Schraw & Dennison, 1994). The literature emphasizes that metacognition is closely connected to various cognitive domains, including academic achievement, problem-solving, critical thinking, and self-regulation (Veenman et al., 2006). Therefore, supporting students' metacognitive skills in educational settings enables them to take responsibility for their own learning and achieve lasting comprehension.

In this study, a metacognitive awareness scale was developed, and its validity and reliability were

examined better to understand students' awareness of their learning processes, their ways of thinking, and how they monitor and evaluate their own learning. Students' awareness of what and how they learn is closely related to their ability to plan, monitor, and regulate the learning process. Accordingly, the developed scale aims to reveal how students apply their metacognitive skills in different learning contexts across various subjects.

This research also aims to contribute to the development of more conscious and practical approaches among educators and future researchers in understanding students' thinking processes. The validity and reliability study of the scale was conducted with 1,009 students enrolled in the 5th, 6th, 7th, and 8th grades at four different middle schools located on the European side of Istanbul during the first semester of the 2025–2026 academic year.

To evaluate the content and face validity of the items in the initial pool, expert opinions were gathered from two professors and two associate professors who have conducted research in the field of metacognition. Additionally, feedback was received from a research assistant specialized in Curriculum and Instruction, an experienced classroom teacher, and a Turkish language teacher responsible for reviewing linguistic clarity and grammar. In total, contributions were provided by six experts and one teacher.

The experts were asked to evaluate each item in the pool as "appropriate," "inappropriate," or "needs revision." Based on their assessments, necessary revisions and modifications were applied to the relevant items. The initial pool, consisting of 60 items, was reviewed by six academicians and one Turkish language teacher to ensure content and face validity. Following the expert evaluations concerning content accuracy, clarity, and conceptual consistency, 14 items were removed from the scale. Specifically, five items were eliminated due to unclear expressions, three because they duplicated other items, three for not adequately representing metacognitive concepts, and three for containing multiple ideas within a single statement. After these revisions, the scale was reduced to 46 items.

Subsequently, to assess the clarity and comprehensibility of the 46-item version, the items were read aloud to a group of 20 students from the 5th, 6th, 7th, and 8th grades. Students were asked to identify any statements they found difficult to understand. Based on their feedback, necessary adjustments were applied. It was noted that 5th-grade students had particular difficulty understanding abstract terms such as "method," "strategy," and "mnemonic." To address this issue, these terms were accompanied by parenthetical examples to enhance comprehension.

As a result of these procedures, the face and content validity of the scale were confirmed, and the instrument was finalized for implementation. The Metacognitive Awareness Scale for Middle School Students was administered during the first semester of the 2025–2026 academic year to a total of 1,009 students (grades 5–8) attending four different middle schools in Istanbul. The validity and reliability analyses of the scale were conducted using data collected from this administration. Specifically, exploratory factor analysis (EFA) was conducted on data from 475 students, and confirmatory factor analysis (CFA) was performed on data from 534 students.

Data Analysis

The Metacognitive Awareness Scale for Middle School Students, developed within the scope of this study, was administered to the designated sample group, and the collected data were analyzed to examine the scale's psychometric properties. The analysis was conducted using SPSS 22.0 statistical software. To evaluate the normality of data distribution, skewness and kurtosis values were examined. According to George and Mallery (2010), values between -1 and +1 indicate that the data are normally distributed. The analysis revealed a skewness value of 0.087 and a kurtosis value of 0.787, confirming that the dataset exhibited a normal distribution.

Within the scope of validity analyses, both content validity and construct validity of the scale were examined. Content validity was ensured through expert opinion, as detailed in the section titled Development of Data Collection Instruments. To determine construct validity, Exploratory Factor Analysis (EFA) was conducted on the first independent sample group ($n = 475$). The analysis was performed using the Principal Axis Factoring (PAF) method, which is more appropriate for identifying latent constructs as it focuses on shared variance among items rather than total variance. Given that metacognitive components are theoretically related, an oblique rotation method (Direct Oblimin) was applied. To determine the number of factors, multiple criteria were considered, including eigenvalues, scree plot, and parallel analysis. The results of the parallel analysis supported a single-factor structure. The number of factors to be retained was determined based on the Kaiser criterion (eigenvalues > 1) and a rigorous examination of the Scree Plot. The scree plot clearly indicated a single-factor structure, showing a sharp break after the first component. Before conducting the factor analysis, the Kaiser-Meyer-Olkin (KMO) coefficient and Bartlett's Test of Sphericity were calculated to assess the suitability of the data. The KMO value of 0.952 and significant Bartlett's test result ($\chi^2 = 8377.999$, $p < .01$) confirmed that the dataset was highly suitable for factor analysis. To confirm the factor structure obtained from EFA,

Confirmatory Factor Analysis (CFA) was subsequently performed with a second independent sample group ($n = 534$). In line with established guidelines, model fit was evaluated using multiple indices, including χ^2/df , CFI, TLI, IFI, and RMSEA.

Additionally, the reliability of the scale was assessed using internal consistency analysis. The analysis yielded a Cronbach's alpha coefficient of .942, indicating a very high level of internal consistency. This result suggests that the items of the scale consistently measure the same construct. Furthermore, test-retest reliability was assessed to evaluate the temporal stability of the scale. The Intraclass Correlation Coefficient (ICC) value was found to be 0.997, demonstrating excellent consistency. This finding indicates that the scale produces stable and dependable results when administered at different times. To provide additional evidence for the validity of the instrument, item-total correlation analysis was also performed. The correlation coefficients ranged from .467 to .651, indicating that all items were significantly and moderately to strongly correlated with the total scale score. This finding confirms that each item corresponds with the overall structure of the scale and contributes meaningfully to the construct being measured.

Findings

Results Related to Construct Validity

Exploratory Factor Analysis

In this study, the sample size was first examined to determine the appropriateness of the data for factor analysis. Although different views exist in the literature regarding sample size, samples smaller than 30 are generally considered weak, while samples up to 500 are regarded as sufficient for most analyses (Büyüköztürk et al., 2009). To further evaluate the data's appropriateness for factor analysis, the Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy and Bartlett's Test of Sphericity were conducted (Barnett & Kaiser, 1978). A KMO value of 0.50 or above indicates that the data are suitable for factor analysis, while a significant Bartlett's test ($p < .05$) suggests sufficient inter-variable correlations (Tavşancıl, 2019; Williams et al., 2010). The analysis revealed a KMO value of 0.952 and a significant Bartlett's test result ($\chi^2 = 8377.999$, $p < 0.01$), confirming that the dataset was highly suitable for factor analysis.

During the Exploratory Factor Analysis (EFA), factor loadings were reviewed to determine whether the items represented a specific construct. A factor loading value of .40 or higher is generally accepted as sufficient (Büyüköztürk, 2002). It is recommended that an item's loading on its primary factor be substantially higher than its loading on other factors, with a minimum difference of .10. Items with low

communality values ($< .30$) are considered to have weak representation within the construct (Costello & Osborne, 2005). However, theoretical consistency was also taken into account alongside statistical criteria during item evaluation (Rattray & Jones, 2007).

In line with these principles, the initial 46-item form was analyzed, and 17 items were removed based on statistical and theoretical considerations. The reasons for item removal are summarized below:

- Items 2, 5, 6, 8, 13, 42, 44, 45, and 46 were removed because their factor loadings were below .40.
- Items 16, 23, 24, and 26 were excluded due to high loadings on multiple factors (loading difference $< .10$), which could reduce conceptual consistency.
- Items 27, 29, 33, and 39 were removed because of low communality values ($< .30$), indicating poor representation of the underlying construct.

Following item removal, Exploratory Factor Analysis using the Principal Axis Factoring (PAF) method with Direct Oblimin rotation was conducted. The number of factors was determined using parallel analysis as the primary and most robust criterion, supported by eigenvalues and the scree plot. All criteria consistently

indicated a single-factor structure.

As a result, a single-factor structure consisting of 29 items was obtained, explaining 38.392% of the total variance. In addition, the results of the parallel analysis provided further support for the unidimensional structure, confirming that a single-factor solution was the most appropriate representation of the data. Although metacognition is theoretically conceptualized as a multidimensional construct, the findings of the present study revealed a unidimensional structure. This result can be explained by the developmental characteristics of middle school students, for whom metacognitive processes may function as an integrated and holistic system rather than as clearly differentiated components.

These findings indicate that the scale represents metacognitive awareness as a unified construct among middle school students. Therefore, construct validity was confirmed, and the final version of the scale consisted of 29 items. To further support the validity and reliability of the obtained structure, detailed results of the factor analysis are presented in Table 2, which includes factor loadings item means, standard deviations, item-total correlation coefficients, and other relevant data.

Table 2.

Results of the Factor Analysis for the Metacognitive Awareness Scale for Middle School Students

Item No	Item Statement (English Translation)	M	SD	Item-Total Correlation	Factor Loading
m1	I know which sources (books, videos, lecture notes, etc.) I should use to learn a topic well.	4.07	1.12	.649	.688
m3	I notice the mistakes I make while studying.	3.72	1.07	.551	.589
m4	When I do not understand a topic, I try to learn it by using different methods (reading another source, asking the teacher, trying again, etc.).	3.95	1.13	.597	.633
m7	I know which subjects are easy and which are difficult for me.	3.94	1.23	.556	.596
m9	Before starting a new topic, I think about what I already know and what I need to learn.	3.57	1.21	.511	.543
m10	While learning something, I try to understand why that information is important.	3.76	1.18	.561	.595
m11	I think about how I can use what I have learned in real life.	3.69	1.25	.535	.568
m12	While learning, I ask myself, "How can I understand this topic better?"	3.72	1.14	.571	.603
m14	To avoid forgetting important information, I create reminders (taking notes, setting an alarm on my phone, etc.).	3.64	1.36	.467	.501
m15	Visuals, graphs, and tables help me understand the topic I am studying.	3.61	1.25	.488	.519
m17	Before studying, I determine which topics I need to learn.	3.77	1.23	.625	.661
m18	I plan in advance how to use my time while studying.	3.64	1.23	.508	.542
m19	I prepare the materials I will use (books, videos, notes, etc.) beforehand.	3.82	1.23	.603	.638
m20	I organize a suitable study environment (quiet place, tidy desk, etc.) to concentrate better.	4.00	1.22	.595	.635
m21	While studying, I decide which information to learn first and which to learn later.	3.55	1.27	.626	.662
m22	I know that I need to spend more time on topics that are difficult to understand.	3.84	1.13	.620	.658
m25	I check whether I understand the topic while studying.	3.88	1.18	.651	.688
m28	I assess myself to check what I have learned.	3.81	1.10	.591	.627
m30	I try to identify which parts of a topic are difficult for me.	3.87	1.13	.597	.636
m31	I underline or take notes on the important information in the text I read.	3.86	1.19	.614	.652
m32	When solving questions, I think about which information I need.	3.76	1.17	.547	.583
m34	While studying a topic, I check how much progress I have made.	3.76	1.13	.617	.655
m35	I make sure to complete the assignments given by my teacher on time.	3.97	1.13	.646	.685
m36	While learning a topic, I try to determine which information is most important for me.	3.72	1.16	.618	.655
m37	I reuse the learning methods that worked well before (reviewing my notes, practicing, using visuals, etc.).	3.93	1.13	.640	.678
m38	I review the methods I use while studying (summarizing, listening, solving questions, practicing, using visuals, etc.) and change them if necessary.	3.72	1.17	.550	.588
m40	I organize my study environment to improve my concentration.	3.94	1.18	.593	.629
m41	While studying, I decide which learning method (writing, listening, watching, etc.) helps me learn better.	3.90	1.14	.599	.636
m43	After an exam, I think about how I can correct the questions I got wrong.	3.83	1.18	.531	.567

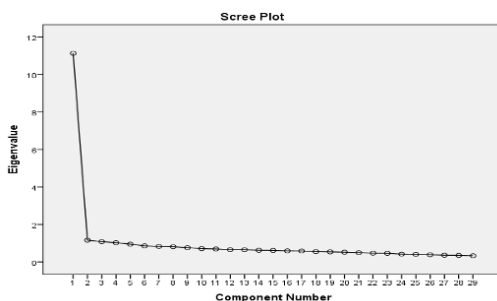
The correlation analysis conducted among the items of the scale revealed that all items were positively and statistically significantly correlated with one another ($p < .01$). This finding indicates that the items were moderately to highly interrelated. Examination of the table demonstrates that all items have significant ($p < .01$) and moderate to strong correlations with the total scale score.

These correlation values confirm that the items of the scale are valid and measure the same theoretical construct. The adequacy of the item-total correlation coefficients provides important evidence for the internal consistency and homogeneity of the scale. The examination of both item-item and item-total correlations suggest that the statements included in the scale sufficiently represent the metacognitive construct intended to be assessed.

To provide a more comprehensive evaluation of the factor structure of the scale, a scree plot illustrating the factor loadings was created (Figure 2). This visual representation supports the analytical findings and clearly assists in determining the number of factors.

Figure 2.

Scree Plot of Eigenvalues for the Metacognitive Awareness Scale for Middle School Students



Confirmatory Factor Analysis

Confirmatory Factor Analysis (CFA) is a statistical method used to test whether the data obtained from a sample fit a theoretically specified model (Sun, 2005). To evaluate model fit, several fit indices are examined. Since each index has its own strengths and limitations, it is recommended to use multiple indices when assessing model fit (Cole, 1987; Sümer, 2000). Commonly used indices include the Chi-square goodness-of-fit test (χ^2), Goodness of Fit Index (GFI), Adjusted Goodness of Fit Index (AGFI), Comparative Fit Index (CFI), Incremental Fit Index (IFI), Tucker-Lewis Index (TLI), Root Mean Square Residual (RMR), and Root Mean Square Error of Approximation (RMSEA).

In this study, a CFA was performed on the 29-item single-factor model identified through Exploratory Factor Analysis (EFA) to confirm the validity of the obtained factor structure. The CFA results demonstrated that all items loaded on a single factor. This represents the

overall structure of metacognition. In other words, all items in the scale were found to contribute to measuring metacognition as a unified construct, confirming that the scale reflects a unidimensional model of metacognitive awareness. Detailed CFA results are presented in Table 3.

Table 3.

Confirmatory Factor Analysis Results for the Metacognitive Awareness Scale for Middle School Students

Model	χ^2/df	NFI	RFI	IFI	TLI	CFI	RMSEA
Metacognitive Awareness Scale for Middle School Students	2.770	.870	.860	.904	.899	.904	.048

When evaluating the fit indices obtained from the CFA, the χ^2/df ratio was found to be 2.770, which falls within the acceptable range reported in the literature and indicates that the model fits the data adequately (Schermelleh-Engel et al., 2003). The other fit indices were calculated as follows: NFI = .87, RFI = .86, IFI = .904, TLI = .899, CFI = .904, and RMSEA = .048. These findings demonstrate that the model tested in the CFA presents an acceptable level of fit with the observed data.

Sample size is an important factor in interpreting the results of confirmatory factor analysis; however, the literature presents differing perspectives on this issue. Some researchers argue that there should be at least 10 to 20 participants per item, whereas others suggest that the sample size should be at least 200 participants (Hox & Bechger, 1998; Rosseel, 2020). The sample size used in the present study met at least one of these criteria, supporting the validity of the analyses.

The path diagram of the confirmatory factor analysis model is presented in Figure 3, while the regression coefficients related to the model are detailed in Table 4. All relationships among the items within the model were found to be statistically significant ($p < .05$), indicating strong evidence for the scale's structural validity.

Figure 3.

Path Diagram of the Confirmatory Factor Analysis for the Metacognitive Awareness Scale for Middle School Students

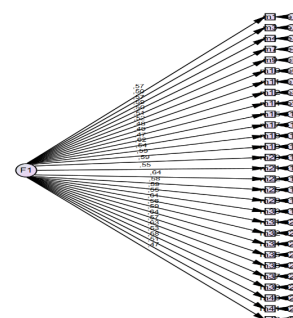


Table 4.
Regression Analysis Results for the Metacognitive Awareness Scale for Middle School Students

Item	Factor	Factor Loading	Standard Error	C.R.	p
M1	Factor 1	.572	—	—	—
M3	Factor 1	.502	.059	13.636	***
M4	Factor 1	.574	.063	15.109	***
M7	Factor 1	.594	.065	15.480	***
M9	Factor 1	.504	.064	13.682	***
M10	Factor 1	.508	.066	13.755	***
M11	Factor 1	.521	.067	14.039	***
M12	Factor 1	.482	.065	13.213	***
M14	Factor 1	.492	.068	13.417	***
M15	Factor 1	.473	.064	13.009	***
M17	Factor 1	.621	.069	15.991	***
M18	Factor 1	.542	.066	14.468	***
M19	Factor 1	.592	.064	15.454	***
M20	Factor 1	.585	.067	15.319	***
M21	Factor 1	.548	.069	14.588	***
M22	Factor 1	.644	.066	16.396	***
M25	Factor 1	.581	.066	15.230	***
M28	Factor 1	.590	.063	15.403	***
M30	Factor 1	.552	.066	14.668	***
M31	Factor 1	.606	.064	15.713	***
M32	Factor 1	.556	.064	14.739	***
M34	Factor 1	.589	.064	15.398	***
M35	Factor 1	.637	.064	16.266	***
M36	Factor 1	.571	.068	15.041	***
M37	Factor 1	.627	.066	16.100	***
M38	Factor 1	.525	.065	14.120	***
M40	Factor 1	.575	.064	15.128	***
M41	Factor 1	.513	.065	13.878	***
M43	Factor 1	.469	.062	12.912	***

Note. p < .001

As shown in Table 4, the factor loadings of the items clustered under a single factor range from .469 (M43) to .644 (M22). All factor loadings were found to be statistically significant ($p < .001$). These results indicate that each item sufficiently represents the factor it belongs to and that the overall structure of the scale aligns well with its intended purpose.

Reliability Findings

Internal consistency is a crucial indicator of reliability that demonstrates the extent to which the items within a measurement instrument represent the same construct and evaluate the internal coherence among them. It expresses the overall integrity of the test and the association of each item with the total score. The internal consistency coefficient generally ranges between 0 and 1, with higher values indicating greater similarity among the items of the scale (Bonett & Wright, 2015; Tavakol & Dennick, 2011; Taber, 2018).

The reliability of the instrument was assessed using the Cronbach's alpha (α) coefficient. In the literature, α values between .80 and 1.00 are widely accepted

as indicators of high reliability (Cortina, 1993; Bonett, 2002). Following the factor analysis conducted within the scope of the research, the 29-item scale yielded a Cronbach's alpha coefficient of .942, indicating that it is a highly reliable measurement tool. Although the Cronbach's alpha coefficient (.942) is relatively high and may suggest potential item redundancy, the items were carefully designed to represent different aspects of metacognitive processes. Therefore, the high internal consistency is interpreted as an indication of strong coherence among items rather than redundancy. Detailed results of the reliability analysis are presented in Table 5.

Table 5.
Reliability Analysis Results for the Metacognitive Awareness Scale for Middle School Students

Scale	Cronbach's Alpha (α)
Metacognitive Awareness Scale for Middle School Students	.942

To determine the stability of the scale scores over time, a test-retest reliability analysis was conducted (DeVellis, 2017). This method involves administering the same measurement instrument to the same group of participants at two different points in time and examining the correlation between the scores obtained from both administrations. It is recommended that the interval between administrations be determined based on the nature of the construct being measured (Streiner et al., 2015). The literature suggests that a time gap of two to four weeks is appropriate and that data from at least 30 participants are sufficient for this type of analysis (Tavşancıl, 2019; Seğer, 2018).

Accordingly, to assess the temporal stability of the scale, two administrations conducted three weeks apart were compared using the Intraclass Correlation Coefficient (ICC). The analysis yielded an ICC(3,1) = 0.997, with a 95% confidence interval [0.996, 0.998]. Based on the classification proposed by Koo and Li (2016), ICC values below 0.50 indicate poor reliability, values between 0.50 and 0.75 indicate moderate reliability, values between 0.75 and 0.90 indicate good reliability, and values above 0.90 indicate excellent reliability. Therefore, the obtained result demonstrates that the developed scale provides highly stable and precise measurements over time.

Scoring of the Scale

In Likert-type scales, an individual's total score is typically calculated by summing or averaging the responses provided to each item (Vaske et al., 2017). Such rating instruments are also known as summated rating scales. However, several studies have suggested that the inclusion of reverse-coded (negatively worded) items may negatively impact the internal consistency of the scale and lead to systematic

measurement errors in the factor structure (Cheng & Hamid, 1997; Turgut & Erden, 2013; Wong et al., 2003). For this reason, no negatively worded items were incorporated in the developed scale.

During the analysis process, it was observed that some initially designed reverse items did not load on the expected factor and reduced the internal consistency coefficient. Since these items failed to establish meaningful relationships with the overall construct, they were removed from the scale. Consequently, the final version of the scale consists of 29 positively worded items only. Each item on the scale is rated using a five-point Likert-type format, ranging from 1 (Strongly Disagree) to 5 (Strongly Agree). Accordingly, respondents indicate their degree of agreement with each statement on a continuum where 1 represents complete disagreement, 2 indicates disagreement, 3 reflects neutrality, 4 denotes agreement, and 5 signifies complete agreement. This structure allows for the measurement of participants' metacognitive awareness levels based on the extent to which they endorse each item. The lowest possible total score on the scale is 29, and the highest possible score is 145. Higher scores indicate a greater level of metacognitive awareness among middle school students.

Discussion And Conclusion

During the development of a scale to assess students' metacognitive awareness levels, a comprehensive literature review was conducted, and an initial pool of 60 items was generated. To ensure content and face validity, the items were evaluated by six field experts specializing in metacognition, and a Turkish language teacher assessed the linguistic clarity and grammatical accuracy of the items. Based on expert feedback, 14 items that were either conceptually inadequate or linguistically unclear were removed from the scale.

To test the applicability and reliability of the developed instrument, the scale was administered to 1,009 students enrolled in the 5th, 6th, 7th, and 8th grades at four different middle schools situated on the European side of Istanbul during the first semester of the 2025–2026 academic year. The large sample size, along with data collected from diverse school types and grade levels, enhanced the reliability and generalizability of the results.

Following data collection, an Exploratory Factor Analysis (EFA) was conducted on the initial 46-item version to assess the structural validity of the scale. During the EFA, items with low communalities and factors with fewer than three items were considered unstable and subsequently removed. Using the Principal Axis Factoring (PAF) method with Direct Oblimin rotation, factor loadings, communality values, and conceptual coherence were examined simultaneously. As a result,

17 items (items 2, 5, 6, 8, 13, 16, 23, 24, 26, 27, 29, 33, 39, 42, 44, 45, and 46) were eliminated due to low loadings, cross-loadings, or weak communalities. Ultimately, a single-factor structure comprising 29 items was obtained, explaining 38.392% of the total variance.

Although metacognition is theoretically conceptualized as a multidimensional construct, the unidimensional structure obtained in this study can be explained by the developmental characteristics of middle school students. At this stage, metacognitive processes may operate as an integrated and holistic system rather than as clearly differentiated components.

Although the total variance explained (38.392%) may appear relatively modest, it is considered acceptable in social sciences, particularly for complex and multifaceted constructs such as metacognition. In scale development studies, an explained variance rate of approximately 30% or higher is generally regarded as sufficient to represent the underlying construct (Hair, J. F. et al., 2019; Tabachnick, B. G. & Fidell, L. S., 2013). Previous studies have also reported similar variance levels, supporting the adequacy of this value in representing complex psychological structures (Field, A., 2024).

The structure derived from the EFA was subsequently tested through Confirmatory Factor Analysis (CFA) to verify its validity. The CFA results indicated that the proposed model demonstrated an acceptable fit to the data and was consistent with the theoretical framework. These findings provide strong evidence for the structural validity of the scale and support the adequacy of the unidimensional model. Overall, the results suggest that the scale measures a coherent and valid construct capable of representing middle school students' metacognitive awareness.

To evaluate reliability, the Cronbach's alpha coefficient was calculated, yielding a value of .942, indicating a highly reliable measurement instrument. In general, coefficients between $.60 \leq \alpha < .80$ suggest adequate reliability, whereas coefficients between $.80 \leq \alpha \leq 1.00$ indicate strong reliability. Accordingly, the developed Metacognitive Awareness Scale for Middle School Students demonstrates strong internal consistency.

Additionally, test-retest reliability was examined to assess temporal stability. The Intraclass Correlation Coefficient (ICC) calculated from two administrations conducted three weeks apart was $ICC(3,1) = .997$, 95% CI [0.996, 0.998]. According to Koo and Li (2016), ICC values above .90 indicate excellent reliability. This result confirms that the scale produces stable and consistent measurements over time.

The findings of the present study are consistent with previous research on metacognitive awareness while

also extending the existing literature by providing evidence from a developmentally appropriate instrument specifically designed for middle school students. For example, Karakelle and Saraç (2007) reported that the Children's Metacognitive Awareness Scale (CMAS-C), originally developed with a multifactorial structure for students in Grades 3–9, was subsequently interpreted as a unidimensional construct through the use of total scores. Similarly, Çetinkaya and Erkin (2002) developed a metacognitive inventory for sixth-grade students consisting of four dimensions: evaluation, self-control, awareness, and cognitive strategies. Akin et al. (2007) validated the Turkish version of Schraw and Dennison's (1994) Metacognitive Awareness Inventory (MAI) for university students. Their findings supported an eight-factor structure comprising declarative, procedural, and conditional knowledge together with planning, monitoring, information management, debugging, and evaluation. Taken together, these findings suggest that although metacognitive processes are theoretically conceptualized as distinct dimensions, they may operate as an integrated system during middle school, supporting the unidimensional structure identified in the present study.

Yıldız et al. (2009) developed a metacognition scale for students in Grades 6–8 consisting of eight dimensions representing both metacognitive knowledge and metacognitive regulation. Whereas their study supported a multidimensional structure, the present study identified a unidimensional structure based on empirical evidence obtained through exploratory and confirmatory factor analyses. Although both instruments address similar metacognitive processes, the present scale conceptualizes these processes as manifestations of a single underlying construct rather than as independent dimensions.

Namlu (2004) developed a four-factor Metacognitive Learning Strategies Scale for university students, focusing primarily on planning, monitoring, and evaluation strategies. In contrast, the present study supports a more parsimonious single-factor model that facilitates score interpretation and practical application in school settings. Similarly, the Motivational, Cognitive, and Metacognitive Competencies Scale adapted by Aktamış and Uça (2010) combines motivational, cognitive, and metacognitive competencies within a broader framework rather than measuring metacognitive awareness independently. Similarly, the Metacognitive Learning Strategies Determination Scale (MLSDDS) by Çöğenli and Güven (2014) targeted preservice teachers and identified four factors—planning, monitoring, evaluation, and affective regulation. Unlike these multidimensional instruments, the present scale conceptualizes these components as indicators of a single latent construct, resulting in a more

parsimonious measurement model that may be more practical for educational assessment. Similarly, Demir (2013) proposed a three-dimensional metacognition scale for preservice teachers, again highlighting that previous instruments have predominantly focused on older populations rather than middle school students.

Considering these studies collectively, the unidimensional Metacognitive Awareness Scale for Middle School Students developed in the present research shows theoretical and structural similarities with both single-factor (e.g., Akin et al., 2007; Demir, 2013; Karakelle & Saraç, 2007) and multidimensional but conceptually integrative scales (e.g., Aktamış & Uça, 2010; Çöğenli & Güven, 2014; Namlu, 2004; Yıldız et al., 2009). The primary contribution of the present study lies in the development of a psychometrically validated instrument specifically designed for students in Grades 5–8, using data collected from diverse school settings and addressing the limited availability of developmentally appropriate metacognitive assessment tools in Türkiye. The strong psychometric properties of the scale, together with its practical applicability, suggest that it can serve as a reliable instrument for both educational research and classroom practice. From a developmental perspective, the unidimensional structure identified in this study may indicate that middle school students do not yet differentiate metacognitive knowledge and regulatory processes as separate cognitive systems. Instead, these processes may operate as a unified construct during early adolescence.

The findings of the present study are consistent with foundational theories of metacognition, which conceptualize metacognition as involving both awareness of cognition and the regulation of cognitive processes (Flavell, 1979; Schraw & Dennison, 1994; Veenman et al., 2006). During middle school, metacognitive awareness and regulatory skills play an important role in improving learning quality, supporting self-regulated learning, and promoting academic development (Kuhn, 2000; Veenman et al., 2006). Previous studies have shown that students with higher levels of metacognition tend to exhibit more stable academic performance, employ more effective learning strategies, and demonstrate greater success in problem-solving. Moreover, metacognition supports self-regulation and independent learning by enabling learners to plan, monitor, and evaluate their own learning processes.

The Metacognitive Awareness Scale for Middle School Students developed in this study differs from earlier instruments both theoretically and in measurement scope. For example, Karakelle and Saraç (2007) primarily emphasized cognitive awareness, whereas Akin et al. (2007) adapted the MAI for university students. Consequently, neither instrument was

specifically designed to reflect the developmental and linguistic characteristics of middle school learners. Similarly, Yıldız et al. (2009) employed an eight-factor scale to differentiate metacognition into several subcomponents. In contrast, the present study supports a unidimensional representation in which cognitive awareness, planning, monitoring, evaluation, and regulation function as interrelated indicators of a single latent construct. This interpretation is supported by both the exploratory and confirmatory factor analyses conducted in the present study.

Unlike previous instruments that primarily emphasized learning strategies (Namlu, 2004; Qögenli & Güven, 2014), the present scale was specifically designed to assess metacognitive awareness as a distinct psychological construct. The Motivational, Cognitive, and Metacognitive Competencies Scale (Aktamış & Uça, 2010), though comprehensive, did not isolate metacognition as a distinct construct. In contrast, the current scale evaluates students' metacognitive awareness through observable learning behaviours, measuring not only "knowing what one knows" but also the ability to plan, monitor, regulate, and evaluate one's own learning. Accordingly, the scale reflects the practical dimension of metacognitive awareness by capturing how students consciously manage their learning processes. Moreover, it provides a developmentally appropriate and educationally relevant measure for assessing metacognitive awareness among middle school students.

The primary contribution of this study lies in establishing a developmentally appropriate, process-oriented, and culturally relevant measurement tool for assessing metacognitive awareness in middle school students. Unlike existing instruments, the present scale integrates both metacognitive knowledge and regulatory processes within a unified structure, reflecting the holistic nature of metacognitive functioning in this age group. In this respect, the present study addresses a critical gap in the literature by offering a scale specifically designed for middle school students, both in terms of developmental appropriateness and alignment with contemporary educational practices.

From a practical perspective, the developed scale provides teachers with a valid and reliable tool for monitoring students' metacognitive awareness, identifying areas requiring instructional support, and designing targeted learning interventions. In addition, school counselors and educational psychologists may employ the instrument to evaluate students' metacognitive development, while researchers may use it to investigate the relationships between metacognitive awareness and variables such as academic achievement, motivation, self-regulated learning, and critical thinking. Moreover, the scale can be used in intervention studies to evaluate the

effectiveness of instructional programs aimed at improving metacognitive skills.

Limitations and Future Research

Although the present study provides strong evidence for the validity and reliability of the Metacognitive Awareness Scale for Middle School Students, several limitations should be acknowledged. First, the sample was restricted to middle schools located in specific districts of Istanbul, which may limit the generalizability of the findings to other regions and socio-cultural contexts. Future studies should therefore replicate the scale development process with more diverse samples across different geographical and cultural settings. Second, the data relied exclusively on self-report measures, which may be subject to social desirability bias and students' limited ability to accurately reflect on their own cognitive processes. Incorporating complementary methods such as teacher observations, performance-based assessments, or mixed-method approaches would provide a more comprehensive evaluation of metacognitive awareness. Third, although metacognition is theoretically conceptualized as a multidimensional construct, the analyses revealed a unidimensional factor structure. This outcome may reflect the developmental characteristics of middle school students, yet future research should explore whether multidimensional models can be identified in different age groups or through alternative measurement approaches. Finally, the study employed a cross-sectional design, which does not allow for examination of developmental changes in metacognitive awareness over time. Longitudinal studies would be valuable in tracking how metacognitive skills evolve throughout middle school and beyond. Addressing these limitations will not only strengthen the psychometric robustness of the scale but also contribute to a deeper understanding of metacognitive development in diverse educational contexts.

Author Contributions: All authors contributed equally to the conception, design, analysis, and writing of this study.

Conflict of Interest: The authors declare that there is no conflict of interest among them.

Ethical Declaration: We declare that the rules specified in the Higher Education Institutions Scientific Research and Publication Ethics Guidelines have been followed and that no actions contrary to scientific research and publication ethics have been taken. All authors contributed to the study, and all responsibility for any potential ethical violations lies with the authors.

Ethics Committee Approval: Ethical approval for this study was obtained from the Social and Humanities

Sciences Ethics Committee of Istanbul University-Cerrahpaşa (Date: 04.03.2025, No: 2025/172).

Funding: This research has not received any funding.

Copyright: The data created or analyzed during this study are available from the authors upon reasonable request.

Data Availability Statement: The data created or analyzed during this study are available from the authors upon reasonable request.

Use of Artificial Intelligence in Writing Assistance: No artificial intelligence tools were used during the writing of this manuscript. However, artificial intelligence-based tools (Napkin AI) were used solely for the creation of Figure 1. The research design, data collection, analysis, and interpretation were carried out entirely by the authors.

References

- Abacı, R., Çetin, B., & Akin, A. (2006). Biliş ötesi farkındalık envanteri: Geçerlik ve güvenirlik çalışması. 14. Ulusal Psikoloji Kongresi Özet Kitapçığı, Hacettepe Üniversitesi, Ankara.
- Akin, A., Abacı, R., & Çetin, B. (2007). The validity and reliability study of the Turkish version of the Metacognitive Awareness Inventory. *Educational Sciences: Theory & Practice*, 7(2), 655–680.
- Aktamış, H., & Uça, S. (2010). Motivasyonel, bilişsel ve bilişüstü yeterlilikler ölçeği'nin Türkçeye uyarlanması. *İlköğretim Online*, 9(3), 980–989.
- Annevirta, T., Laakkonen, E., Kinnunen, R., & Vauras, M. (2007). Developmental dynamics of metacognitive knowledge and text comprehension skill in the first primary school years. *Metacognition and Learning*, 2, 21–39. <https://doi.org/10.1007/s11409-007-9002-8>
- Baker, L., & Brown, A. L. (1980). Metacognitive skills and reading. In P. D. Pearson (Ed.), *Handbook of reading research* (pp. 1–74). Longman.
- Barnett, M., & Kaiser, D. (1978). The relationship between intellectual-achievement responsibility attributions and performance. *Child Study Journal*, 8, 209–215.
- Bonett, D. G. (2002). Sample size requirements for testing and estimating coefficient alpha. *Journal of Educational and Behavioral Statistics*, 27(4), 335–340. <https://doi.org/10.3102/10769986027004335>
- Bonett, D. G., & Wright, T. A. (2015). Cronbach's alpha reliability: Interval estimation, hypothesis testing, and sample size planning. *Journal of Organizational Behavior*, 36(1), 3–15. <https://doi.org/10.1002/job.1960>
- Büyüköztürk, Ş. (2002). Faktör analizi: Temel kavramlar ve ölçek geliştirmede kullanımı. *Kuram ve Uygulamada Eğitim Yönetimi*, 32, 470–483.
- Büyüköztürk, Ş., Kılıç-Çakmak, E., Akgün, Ö. E., Karadeniz, Ş., & Demirel, F. (2009). *Bilimsel araştırma yöntemleri* (3. baskı). Pegem Akademi.
- Cheng, S. T., & Hamid, P. N. (1997). Dispositional optimism in Chinese people: What does the life orientation test measure? *International Journal of Psychology*, 32(4), 221–230.
- Cole, D. A. (1987). Utility of confirmatory factor analysis in test validation research. *Journal of Consulting and Clinical Psychology*, 55(4), 584–594. <https://doi.org/10.1037/0022-006X.55.4.584>
- Conway, J. M., & Huffcutt, A. I. (2003). A review and evaluation of exploratory factor analysis practices in organizational research. *Organizational Research Methods*, 6(2), 147–168. <https://doi.org/10.1177/1094428103251621>
- Cortina, J. M. (1993). What is coefficient alpha? An examination of theory and applications. *Journal of Applied Psychology*, 78(1), 98–104. <https://doi.org/10.1037/0021-9010.78.1.98>
- Costello, A. B., & Osborne, J. W. (2005). Best practices in exploratory factor analysis: Four recommendations for getting the most from your analysis. *Practical Assessment, Research, and Evaluation*, 10(1), 7. <https://doi.org/10.7275/jyj1-4868>
- Coutinho, S. A., & Neuman, G. (2008). A model of metacognition, achievement goal orientation, learning style and self-efficacy. *Learning Environments Research*, 11(2), 131–151. <https://doi.org/10.1007/s10984-008-9043-8>
- Cross, D. R., & Paris, S. G. (1988). Developmental and instructional analyses of children's metacognition and reading comprehension. *Journal of educational psychology*, 80(2), 131–142.
- Çakıroğlu, A. (2007). Üstbiliş. *Türkiye Sosyal Araştırmalar Dergisi*, 11(2), 21–27.
- Çetinkaya, P., & Erkin, E. (2002). Assessment of metacognition and its relationship with reading comprehension, achievement, and aptitude. *Boğaziçi Eğitim Fakültesi Dergisi*, 19(1), 1–11.

- Çögenli, A., & Güven, M. (2014). Bilişüstü öğrenme stratejileri belirleme ölçeğinin geçerlik ve güvenirlik çalışması. *Dicle Üniversitesi Ziya Gökalp Eğitim Fakültesi Dergisi*, 22, 283–297.
- Demir, Ö. (2013). A validation and reliability study of the metacognition scale in Turkey. *Global Journal of Human Social Science*, 13(10), 26–35.
- Desoete, A. (2008). Multi-method assessment of metacognitive skills in elementary school children: How you test is what you get. *Metacognition and Learning*, 3(3), 189–206. <https://doi.org/10.1007/s11409-008-9026-0>
- DeVellis, R. F. (2017). *Scale development: Theory and applications* (4th ed.). SAGE Publications.
- Field, A. (2024). *Discovering statistics using IBM SPSS statistics*. Sage publications limited.
- Flavell, J. H. (1979). Metacognition and cognitive monitoring: A new area of cognitive-developmental inquiry. *American Psychologist*, 34(10), 906–911. <https://doi.org/10.1037/0003-066X.34.10.906>
- Flavell, J. H. (1985). *Cognitive development*. Prentice-Hall.
- Ford, J. K., MacCallum, R. C., & Tait, M. (1986). The application of exploratory factor analysis in applied psychology: A critical review and analysis. *Personnel Psychology*, 39(2), 291–314. <https://doi.org/10.1111/j.1744-6570.1986.tb00583.x>
- George, D., & Mallery, M. (2010). *SPSS for Windows step by step: A simple guide and reference* (10th ed.). Pearson.
- Güner, P., & Erbay, H. N. (2021). Metacognitive skills and problem-solving. *International Journal of Research in Education and Science*, 7(3), 715–734. <https://doi.org/10.46328/ijres.1594>
- Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2019). *Multivariate data analysis*.
- Howard, B. C., McGee, S., Shia, R., & Hong, N. (2000, April). *Metacognitive self-regulation and problem-solving: Expanding the theory base through factor analysis*. [Paper presentation] American Educational Research Association (AERA) Annual Meeting, New Orleans, LA, United states.
- Hox, J. J., & Bechger, T. M. (1998). An introduction to structural equation modeling. *Family Science Review*, 11(4), 354–373.
- Jacobs, J. E., & Paris, S. G. (1987). Children's metacognition about reading: Issues in definition, measurement and instruction. *Educational Psychologist*, 22, 255–278. https://doi.org/10.1207/s15326985ep2203_4
- Jahangard, Z., Soltani, A., & Alinejad, M. (2016). Exploring the relationship between metacognition and attitudes towards science of senior secondary students through a structural equation modeling analysis. *Journal of Baltic Science Education*, 15(3), 340–349. <https://doi.org/10.33225/jbse/16.15.340>
- Karakelle, S., & Saraç, S. (2007). Çocuklar için Üst Bilişsel Farkındalık Ölçeği (ÜBFÖ-Ç) A ve B formları: Geçerlik ve güvenirlik çalışması. *Türk Psikoloji Yazıları*, 10(20), 87–103.
- Karakelle, S., & Saraç, S. (2010). Üst biliş hakkında bir gözden geçirme: Üstbiliş çalışmaları mı yoksa üstbilişsel yaklaşım mı? *Türk Psikoloji Yazıları*, 13(26), 45–60.
- Kline, R. B. (1998). *Principles and practice of structural equation modeling*. Guilford Press.
- Koo, T. K., & Li, M. Y. (2016). A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *Journal of Chiropractic Medicine*, 15(2), 155–163. <https://doi.org/10.1016/j.jcm.2016.02.012>
- Kuhn, D. (2000). Metacognitive development. *Current Directions in Psychological Science*, 9(5), 178–181. <https://doi.org/10.1111/1467-8721.00088>
- Lecce, S., Zocchi, S., Pagnin, A., Palladino, P., & Taumoepeau, M. (2010). Reading minds: The relation between children's mental state knowledge and their metaknowledge about reading. *Child Development*, 81, 1876–1893. <https://doi.org/10.1111/j.1467-8624.2010.01517.x>
- Mokhtari, K., & Reichard, C. A. (2002). Assessing students' metacognitive awareness of reading strategies. *Journal of Educational Psychology*, 94, 249–259. <https://doi.org/10.1037/0022-0663.94.2.249>
- Namlu, A. G. (2004). Bilişötesi öğrenme stratejileri ölçme aracının geliştirilmesi: Geçerlilik ve güvenirlik çalışması. *Anadolu Üniversitesi Sosyal Bilimler Dergisi*, 2, 123–136.
- Owo, W. J., & Ikwut, E. F. (2015). Relationship between metacognition, attitude and academic achievement of secondary school chemistry students in Port Harcourt, Rivers State. *Journal of Research & Method in Education*, 5(6), 6–12.

- Özsoy, G. (2011). An investigation of the relationship between metacognition and mathematics achievement. *Asia Pacific Education Review*, 12, 227–235. <https://doi.org/10.1007/s12564-010-9129-6>
- Öztürk, E. (2012). Okuma stratejileri üstbilişsel farkındalık envanterinin Türkçe formunun geçerlik ve güvenirlik çalışması. *İlköğretim Online*, 11(2), 292–305.
- Pereira-Laird, J. A., & Deane, F. P. (1997). Development and validation of a self-report measure of reading strategy use. *Reading Psychology: An International Quarterly*, 18, 185–235. <https://doi.org/10.1080/0270271970180204>
- Pintrich, P. R. (2002). The role of metacognitive knowledge in learning, teaching, and assessing. *Theory into Practice*, 41(4), 219–225. https://doi.org/10.1207/s15430421tip4104_3
- Rayner, A. (2005, September). *Reflections on context based science teaching: A case study of physics students for physiotherapy* [Poster presentation]. Annual UniServe Science Blended Learning Symposium Proceedings, Sydney.
- Rattray, J., & Jones, M. C. (2007). Essential elements of questionnaire design and development. *Journal of Clinical Nursing*, 16(2), 234–243. <https://doi.org/10.1111/j.1365-2702.2006.01573.x>
- Rosseel, Y. (2020). Small sample solutions for structural equation modeling. In *Small sample size solutions* (pp. 226–238). Routledge. <https://doi.org/10.4324/9780429273872-19>
- Safran, M. (Ed.) (2019). *Dünya eğitime yön verenler* (Cilt 1-2). Millî Eğitim Bakanlığı Yayınları.
- 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.
- Schraw, G. (1998). Promoting general metacognitive awareness. *Instructional Science*, 26(1–2), 113–125. <https://doi.org/10.1023/A:1003092417751>
- Schraw, G., & Dennison, R. S. (1994). Assessing metacognitive awareness. *Contemporary Educational Psychology*, 19, 460–475. <https://doi.org/10.1006/ceps.1994.1033>
- Schraw, G., & Moshman, D. (1995). Metacognitive theories. *Educational Psychology Review*, 7(4), 351–371. <https://doi.org/10.1007/BF02212307>
- Seğer, İ. (2018). *Psikolojik test geliştirme ve uyarılama süreci: SPSS ve LISREL uygulamaları* (2. baskı). Anı Yayıncılık.
- Sperling, R. A., Howard, B. C., Miller, L. A., & Murphy, C. (2002). Measures of children's knowledge and regulation of cognition. *Contemporary Educational Psychology*, 27, 51–79. <https://doi.org/10.1006/ceps.2001.1091>
- Streiner, D. L., Norman, G. R., & Cairney, J. (2015). *Health measurement scales: A practical guide to their development and use* (5th ed.). Oxford University Press. <https://doi.org/10.1093/med/9780199685219.001.0001>
- Sun, J. (2005). Assessing goodness of fit in confirmatory factor analysis. *Measurement and Evaluation in Counseling and Development*, 37(4), 240–256. <https://doi.org/10.1080/07481756.2005.11909764>
- Sümer, N. (2000). Yapısal eşitlik modelleri: Temel kavramlar ve örnek uygulamalar. *Türk Psikoloji Yazıları*, 3(6), 49–74.
- Tabachnick, B. G., & Fidell, L. S. (2007). *Using multivariate statistics* (5th ed.). Pearson Education.
- Tabachnick, B. G., & Fidell, L. S. (2013). *Using multivariate statistics* (6. Baskı). MA: Pearson.
- Taber, K. S. (2018). The use of Cronbach's alpha when developing and reporting research instruments in science education. *Research in Science Education*, 48(6), 1273–1296. <https://doi.org/10.1007/s11165-016-9602-2>
- Tavakol, M., & Dennick, R. (2011). Making sense of Cronbach's alpha. *International Journal of Medical Education*, 2, 53–55. <https://doi.org/10.5116/ijme.4dfb.8dfd>
- Tavşancıl, E. (2019). *Tutumların ölçülmesi ve SPSS ile veri analizi* (6. baskı). Nobel Yayıncılık.
- Turgut, T., & Erden, N. S. (2013). Olumsuz test ifadelerinin iç tutarlılığa ve faktör yapısına etkileri. *Istanbul University Journal of the School of Business Administration*, 42(2).
- Vaske, J. J., Beaman, J., & Sponarski, C. C. (2017). Rethinking internal consistency in Cronbach's alpha. *Leisure Sciences*, 39(2), 163–173. <https://doi.org/10.1080/01490400.2015.1127189>
- Veenman, M. V. J., Van Hout-Wolters, B. H. A. M., & Afflerbach, P. (2006). Metacognition and learning: Conceptual and methodological considerations. *Metacognition and Learning*, 1, 3–14. <https://doi.org/10.1007/s11409-006-6893-0>

- Vrugt, A., & Oort, F. J. (2008). Metacognition, achievement goals, study strategies and academic achievement: Pathways to achievement. *Metacognition and Learning*, 3, 123–146. <https://doi.org/10.1007/s11409-008-9022-4>
- Williams, B., Onsman, A., & Brown, T. (2010). Exploratory factor analysis: A five-step guide for novices. *Australasian Journal of Paramedicine*, 8(3), 1–13. <https://doi.org/10.33151/AJP.8.3.93>
- Wong, N. Y., Rindfleisch, A., & Burroughs, J. E. (2003). Do reverse-worded items confound measures in cross-cultural consumer research? The case of material values scale. *Journal of Consumer Research*, 30(1), 72–91. <https://doi.org/10.1086/374697>
- Yıldız, E., Akpınar, E., Tatar, N., & Ergin, Ö. (2009). İlköğretim öğrencileri için geliştirilen biliş üstü ölçeği'nin açımlayıcı ve doğrulayıcı faktör analizi. *Kuram ve Uygulamada Eğitim Bilimleri*, 9(3), 1573–1604.
- Young, A., & Fry, J. D. (2008). Metacognitive awareness and academic achievement in college students. *Journal of the Scholarship of Teaching and Learning*, 8(2), 1–10.

Reframing the Reasoning Gap: Cognitive-Domain Achievement Profiles and Their Predictors Among Eighth-Grade Students in Türkiye Across TIMSS 2019 and 2023

Başak Barak^a, Betül Barut^b

Received : 13 July 2026
Revised : 2 September 2026
Accepted : 30 September 2026
DOI : 10.26822/iejee.2026.467

^aCorresponding Author: Başak Barak, Anadolu University, Department of Science and Mathematics Education
E-mail: bbarak@anadolu.edu.tr
ORCID: <https://orcid.org/0000-0001-5536-7625>

^b Betül Barut, Anadolu University, Department of Science and Mathematics Education
E-mail: byayan@anadolu.edu.tr
ORCID: <https://orcid.org/0000-0002-6196-9654>

Abstract

Eighth-grade mathematics achievement in Türkiye is often framed around a presumed weakness in higher-order reasoning, yet the relative standing of the TIMSS cognitive domains has rarely been tested inferentially or across time. Using the nationally representative TIMSS 2019 ($N = 4,077$) and 2023 ($N = 4,925$) Grade 8 samples for Türkiye, we profiled achievement across the three cognitive domains, tested how the profile changed between cycles, and modelled each domain as a function of mathematics self-confidence, liking mathematics, home educational resources, and gender. All estimates used the five plausible values with jackknife (JK2) standard errors; domain contrasts are reported with both between-student (d_{av}) and paired (d_z) effect sizes, and student-level results were corroborated with two-level (student-within-school) models. Contrary to the reasoning-deficit narrative, reasoning was not the weakest domain in either cycle: in 2019 it exceeded applying and knowing, and by 2023 the domains converged, with knowing becoming the relative low point. The 2019–2023 gain was concentrated in the applying domain (+18.4 points, $p = .002$). Across all domains and both cycles, self-confidence ($\beta = .42-.45$) and home educational resources ($\beta = .33-.41$) were the dominant correlates, whereas liking mathematics contributed little once self-confidence was controlled (a suppression pattern), and gender differences were small. The predictor structure was near-identical across domains and stable across cycles; multilevel models further showed the home-resources association to be largely compositional (between schools). These findings reframe the Turkish “reasoning gap” and point to school composition and mathematics self-confidence as policy priorities.

Keywords:

TIMSS, mathematics cognitive domains, reasoning, mathematics self-confidence, Türkiye, large-scale assessment.



www.iejee.com
ISSN: 1307-9298

2026 Published by KURA Education & Publishing.
This is an open access article under the CC BY-NC-ND license. (<https://creativecommons.org/licenses/by/4.0/>)

Introduction

In an increasingly technological society, mathematical competence has become a core skill for all learners, equipping individuals to adapt and meet new challenges

as the demands of the world continue to shift (Philpot et al., 2021). Given this importance, mathematics achievement remains a central focus of education systems and is systematically monitored through international assessments. Among these, the Trends in International Mathematics and Science Study (TIMSS), conducted by the International Association for the Evaluation of Educational Achievement (IEA), has assessed the mathematics achievement of eighth-grade students every four years since 1995, providing nearly three decades of trend data that allow comparisons both over time and across countries (von Davier et al., 2024).

Mathematics is characterized not only by a body of concepts and procedures but also by distinctive ways of reasoning through which mathematical knowledge is constructed, justified, and extended (Ball & Bass, 2003). Consistent with this view, mathematical proficiency is widely regarded as multidimensional rather than a single body of content knowledge. Kilpatrick et al. (2001) conceptualize it as comprising interwoven strands of conceptual understanding, procedural fluency, strategic competence, adaptive reasoning, and productive disposition, while Niss (2003) likewise frames the goal of mathematics education in terms of developing mathematical competencies rather than acquiring content knowledge alone. Reflecting this multidimensional perspective, TIMSS conceptualizes mathematics achievement along both content and cognitive dimensions.

These two dimensions are specified in detail in the TIMSS 2023 mathematics framework. The content dimension specifies the subject matter assessed (number, algebra, geometry, and data and probability), whereas the cognitive dimension specifies the thinking processes students draw on when working with that content. Within each content domain, items are developed to assess all three cognitive domains, so that number, algebra, geometry, and data items alike may require knowing, applying, or reasoning; the cognitive dimension therefore captures how students engage with mathematics, independent of the particular topic. It comprises three domains. Knowing refers to students' recall and understanding of mathematical facts, concepts, and procedures. Applying concerns their ability to use this knowledge to solve routine and real-world problems by selecting appropriate strategies and representing situations mathematically. Reasoning involves the logical, systematic thinking required to analyze complex relationships, make inferences, generalize, and justify solutions (Philpot et al., 2021). This design enables researchers to examine students' achievement profiles across knowing, applying, and reasoning, providing a more nuanced picture than overall scores alone. Although the content and cognitive dimensions are conceptually distinct, performance in each cognitive

domain necessarily draws on students' mathematical content knowledge. This organization parallels the revised taxonomy of Anderson and Krathwohl (2001), which likewise conceptualizes learning as the intersection of knowledge and cognitive processes. Accordingly, TIMSS reports not only overall mathematics achievement but also achievement in each cognitive domain separately (Yin & Foy, 2021), making it possible to identify cognitive achievement profiles that may be obscured by overall scores.

Beyond overall achievement levels, a substantial body of research using TIMSS has examined the student characteristics associated with achievement. Across countries and assessment cycles, mathematics self-confidence, home educational resources, liking mathematics, and gender have emerged as the most frequently investigated student-level variables and have consistently been associated with mathematics achievement (Akyüz, 2014; Lee & Stankov, 2018; Yıldırım, 2019).

Among these variables, mathematics self-confidence has consistently emerged as the strongest correlate of mathematics achievement. International syntheses and meta-analyses show that students with higher mathematics self-confidence achieve substantially higher mathematics scores across countries (Çiftçi & Yıldız, 2019; Lee & Stankov, 2018). Similar evidence has been reported for Türkiye. Using TIMSS 2011, 2015, and 2019 data, Coşkun and Karakaya Özyer (2023) found that mathematics self-confidence was the strongest student-level predictor of mathematics achievement across all three assessment cycles, whereas the unique effects of liking mathematics and valuing mathematics were relatively weak or inconsistent. Home educational resources likewise demonstrate a robust positive association with mathematics achievement in Türkiye (Coşkun & Karakaya Özyer, 2023; Ersan & Rodriguez, 2020), a finding that is broadly consistent with other analyses of Turkish TIMSS data (Koyuncu, 2021).

Gender has also received considerable attention in TIMSS research. However, large cross-national studies indicate that gender differences in mathematics achievement are generally small and inconsistent across countries, suggesting that affective and socioeconomic characteristics explain substantially more variation than gender alone (Else-Quest et al., 2010; Ghasemi & Burley, 2019; Reilly et al., 2019; Oberleiter et al., 2023). Recent evidence from Türkiye similarly suggests that gender differences vary across TIMSS cycles, with female students outperforming males in some cycles but not others (Coşkun & Karakaya Özyer, 2023). Collectively, these findings suggest that mathematics self-confidence and home educational resources are the most consistent correlates of mathematics achievement, whereas

the unique contributions of liking mathematics and gender appear to vary across contexts and assessment cycles. Although this body of research has substantially advanced the understanding of mathematics achievement in TIMSS, it has almost exclusively conceptualized achievement as a single composite outcome. Whether these well-established correlates operate uniformly across the knowing, applying, and reasoning domains that constitute mathematics achievement—or whether their relative importance shifts with the cognitive demands of the tasks—therefore remains an open empirical question.

Because the three TIMSS cognitive domains represent qualitatively different mathematical processes, such variation in the relative importance of these correlates is theoretically plausible. Yet despite this conceptual and practical importance, the way achievement is distributed across the cognitive domains has received comparatively little empirical attention. The broader body of research built on international large-scale assessments such as TIMSS has grown rapidly into a mature field (Hernández-Torrano & Courtney, 2021), yet in a mapping of its dominant research themes, achievement disaggregated by the knowing–applying–reasoning distinction does not emerge as a distinct topic. The field’s attention has instead centered on themes such as measurement and testing, educational policy, equity, and student motivation, with the most visible work focusing on cross-national differences in overall achievement (Hernández-Torrano & Courtney, 2021). The process-oriented cognitive dimension has likewise been comparatively neglected within more fine-grained work. Zhu (2023) notes that most research on mathematics learning trajectories concentrates on knowledge content rather than the cognitive domain. Where the cognitive domains are examined directly, analyses have frequently remained at the level of the assessment framework rather than learner performance. Choi et al. (2026), for instance, explicitly acknowledge that their document-based analysis “does not directly examine learner-level performance,” leaving empirical validation against actual achievement data as a task for future research.

Türkiye offers a particularly informative case for taking up this task. With a long record of participation at the eighth-grade level—in 1999, 2007, 2011, 2015, 2019, and 2023—it provides the sustained, learner-level trend data needed to examine cognitive-domain achievement both over time and against actual student performance rather than at the framework level alone. Its recent trajectory makes such an examination especially worthwhile. Across these cycles, Türkiye’s average mathematics scores have risen consistently. In TIMSS 2023, Turkish eighth-graders achieved a mathematics score of 509, well above the international average of 478 and ranking 13th among the 44 participating countries. Notably, this marks the first time Türkiye surpassed the TIMSS scale

centerpoint of 500 at this grade level, with an increase of 13 points over the 2019 cycle. Beneath this overall figure, however, the three cognitive domains reveal an uneven profile. Turkish eighth-graders scored 505 in knowing, 510 in applying, and 511 in reasoning (Ministry of National Education [MoNE], 2024). This profile is noteworthy because Türkiye is one of only eight countries where reasoning is the single highest-scoring cognitive domain. This ordering also runs counter to the difficulty hierarchy ordinarily expected across the domains, in which knowing is the easiest and reasoning the most demanding (Zhu, 2023).

Although the TIMSS 2023 Türkiye Report and the international results document this cognitive-domain profile, these findings remain largely descriptive. The international report does test each domain score against a country’s overall mathematics average, and on this test Türkiye’s knowing, applying, and reasoning scores do not differ significantly from its overall average at the .01 level (von Davier et al., 2024; Ministry of National Education [MoNE], 2024). Yet this leaves the questions most relevant to interpreting Türkiye’s profile unanswered. First, comparing each domain with the overall composite—which itself incorporates that domain—does not establish whether the three domains differ from one another; whether the descriptive ordering in which reasoning ranks highest reflects a genuine relative strength or merely sampling variation therefore remains untested. Second, the reports do not indicate whether the cognitive-domain profile changed significantly between the 2019 and 2023 cycles. Third, they do not show whether student-level characteristics—such as self-confidence, liking mathematics, gender, and home educational resources—are differentially associated with achievement across the three domains. The present study addresses these gaps by providing an inferential analysis of Turkish eighth-graders’ mathematics achievement across the knowing, applying, and reasoning domains using TIMSS 2019 and 2023 data.

The Present Study: Aims and Hypotheses

To this end, the TIMSS 2019 and 2023 Grade 8 data for Türkiye were analyzed in pursuit of three aims: (a) to profile mathematics achievement across the three cognitive domains and quantify their relative standing; (b) to test whether the domain profile changed between 2019 and 2023; and (c) to examine how mathematics self-confidence, liking mathematics, home educational resources, and gender are associated with achievement in each domain, and whether these associations differ across domains or cycles. Because the design is observational, associations are not interpreted causally.

Four hypotheses were advanced. H1 (relative standing): Reasoning achievement will differ from knowing and applying achievement within each cycle; given

competing expectations in the literature, the direction is tested two-sided. H2 (profile change): The cognitive-domain profile will change between 2019 and 2023. H3 (dominant correlates): Mathematics self-confidence and home educational resources will be the strongest positive correlates of achievement in every domain, whereas liking mathematics will show weak unique associations once self-confidence is controlled. H4 (structural invariance): The configuration of predictors will be largely invariant across the three cognitive domains.

Method

Data and Sample

TIMSS 2019 and 2023 Grade 8 assessments for Türkiye, obtained from the IEA international database were used in the analysis (Mullis et al., 2020). TIMSS uses a two-stage stratified cluster sample in which schools are selected first and intact classrooms within schools second. The Türkiye main samples comprised 4,077 students in 181 schools (2019) and 4,925 students in 141 schools (2023). Only the main (digital) samples were analyzed; the paper bridge and problem-solving and inquiry sub-samples were excluded.

Measures

Outcomes

Mathematics achievement in each cognitive domain—knowing (BSMKNO), applying (BSMAPP), and reasoning (BSMREA)—was represented by five plausible values reported on the TIMSS scale (international centerpoint = 500, $SD = 100$). Plausible values are multiple imputations of proficiency that, combined appropriately, yield unbiased population estimates and correct standard errors (Arikan et al., 2020; Braun & von Davier, 2017).

Predictors

Three IEA-constructed IRT context scales were used: mathematics self-confidence (Students Confident in Mathematics, BSBGSCM), liking mathematics (Students Like Learning Mathematics, BSBGSLM), and home educational resources (BSBGHER); higher scores indicate more of the named attribute. Gender was coded 1 = girl, 0 = boy (ITSEX). Students Confident in Mathematics comprised eight Likert items (e.g., “I usually do well in mathematics”; Cronbach’s $\alpha = .89$ in the Turkish 2023 sample), and Students Like Learning Mathematics comprised nine items (e.g., “I enjoy learning mathematics”; $\alpha = .93$). Home Educational Resources is an IEA-scaled composite index combining the number of books in the home, home study supports (own room and internet access), and parents’ highest education level, and is therefore reported as an index rather than an internal-consistency scale

Analytic Strategy

Analyses were conducted in R 4.5.2 (R Core Team, 2025) using the five plausible values and the total student weight (TOTWGT). Standard errors were estimated with the jackknife repeated-replication procedure (JK2) using the 75 sampling zones and the jackknife replicate indicator, and plausible-value estimates were combined by Rubin’s (1987) rules (Foy & LaRoche, 2020); this implementation reproduced the published Türkiye mathematics means and standard errors, confirming its accuracy. For each cycle we estimated weighted domain means and standardized pairwise domain differences. Two effect sizes are reported for the domain contrasts: d_{av} , standardized by the pooled between-student standard deviation, and—because the domains are strongly correlated within students—the paired effect size d_z , standardized by the standard deviation of the within-student difference scores (Cohen, 1988). Cross-cycle contrasts were tested with independent-samples z tests based on the two cycles’ jackknifed standard errors. Predictors of each domain were estimated with weighted linear regressions (predictors standardized), fitted separately by domain and cycle and pooled across the five plausible values; coefficients are interpreted as associations. Missing data on predictors were limited ($< 3.5\%$ per variable) and handled by listwise deletion.

Results

Following the analytic approach described in Section 2.3, all estimates are weighted (TOTWGT) and pool the five plausible values, and are reported separately for TIMSS 2019 ($N = 4,077$) and 2023 ($N = 4,925$). The three cognitive domains—knowing, applying, and reasoning—are examined first as an achievement profile (weighted domain means and standardized pairwise differences) and then as outcomes in parallel regression models predicted by mathematics self-confidence, liking mathematics, home educational resources, and gender.

Descriptive Statistics and the Cognitive-Domain Profile

Table 1 reports weighted descriptive statistics for all study variables in each cycle, and Figure 1 displays the cognitive-domain achievement profile. In 2019, reasoning ($M = 503.9$) stood above both knowing ($M = 494.1$) and applying ($M = 491.3$); by 2023 all three domains had risen and clustered more closely together (knowing = 504.7, applying = 509.8, reasoning = 511.4), each at or above the TIMSS scale centerpoint of 500. Thus, contrary to the common expectation of a reasoning shortfall, Grade 8 students in Türkiye did not underperform in reasoning relative to the more routine domains in either cycle; instead, the domain that lagged shifted over time.

Table 1

Weighted Descriptive Statistics for Study Variables by TIMSS Cycle

Variable	M 2019	SD 2019	M 2023	SD 2023
Reasoning achievement	503.9	110.6	511.4	111.8
Applying achievement	491.3	103.4	509.8	111.4
Knowing achievement	494.1	126.0	504.7	120.6
Self-confidence (BSBGSCM)	9.84	2.37	9.65	2.34
Liking mathematics (BSBGSLM)	10.32	1.93	10.22	1.97
Home resources (BSBGHER)	9.47	1.83	10.23	1.66
Female (%)	50.2	—	51.4	—
N	4,077		4,925	

Note. Achievement domains are on the TIMSS reporting scale (international centerpoint = 500); context measures are IEA IRT scales. All estimates are weighted (TOTWGT). 2019 N = 4,077; 2023 N = 4,925.

Table 2

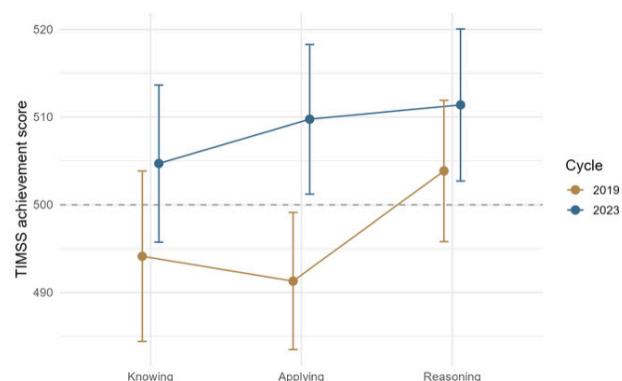
Weighted Bivariate Correlations Among Study Variables

Variable	1	2	3	4	5	6	7
1. Reasoning	—	.96	.96	.44	.25	.47	.06
2. Knowing	.95	—	.96	.47	.27	.48	.07
3. Applying	.97	.96	—	.45	.25	.47	.03
4. Self-confidence	.49	.50	.48	—	.70	.14	-.03
5. Liking mathematics	.31	.31	.30	.67	—	-.01	.02
6. Home resources	.43	.43	.43	.22	.08	—	.05
7. Female	-.01	.02	-.01	-.07	-.04	.00	—

Note. Weighted correlations averaged across five plausible values for achievement domains. Values above the diagonal are TIMSS 2019; below the diagonal are TIMSS 2023. Female coded 1 = girl.

Figure 1

Mathematics Cognitive-Domain Achievement Profile, Türkiye Grade 8



Note. Weighted means with 95% confidence intervals (five plausible values, JK2 jackknife). The dashed line marks the TIMSS scale centerpoint (500).

Pairwise Domain Differences and Cross-Cycle Change

In 2019, reasoning significantly exceeded applying by 12.5 score points ($p < .001$, $d_{av} = 0.12$) and knowing by 9.7 points ($p < .001$, $d_{av} = 0.08$), whereas applying and knowing were practically indistinguishable. By 2023 the reasoning–applying difference had shrunk to

a non-significant 1.6 points ($p = .17$, $d_{av} = 0.01$), and knowing had become the lowest-performing domain: applying exceeded knowing by 5.1 points ($p < .001$) and reasoning exceeded knowing by 6.7 points ($p < .001$). Across cycles (Table 3), the overall improvement was driven primarily by applying, which rose 18.4 points ($p = .002$, $d = 0.17$); the gains in knowing (+10.6) and reasoning (+7.5) were positive but not statistically significant. Consistent with this, the reasoning advantage over applying narrowed significantly ($\Delta = -10.9$, $p < .001$), and applying moved from slightly below to significantly above knowing ($\Delta = +7.9$, $p < .001$; Table 4). All within-cycle domain differences were small in effect-size terms ($|d_{av}| \leq 0.12$), indicating that the domains are close in absolute standing even where differences are statistically reliable. Because the domains were highly correlated within students ($r_s \approx .96$), paired effect sizes exceeded the between-student d_{av} ; for example, the 2019 reasoning-over-applying advantage corresponded to $d_{av} = 0.12$ but $d_z = 0.42$, declining to $d_z = 0.06$ by 2023 (Table 3).

Table 3

Cognitive-Domain Achievement and Cross-Cycle Change (2019 to 2023)

Domain	M 2019 (SE)	M 2023 (SE)	Δ	z	p	d
Knowing	494.1 (4.95)	504.7 (4.57)	+10.6	1.57	.117	0.086
Applying	491.3 (3.99)	509.8 (4.35)	+18.4	3.12	.002	0.172
Reasoning	503.9 (4.11)	511.4 (4.42)	+7.5	1.25	.212	0.068

Note. Weighted means with jackknife standard errors in parentheses. d = Cohen's d for the cross-cycle change. Cross-cycle tests are independent-samples z tests.

Table 4

Pairwise Cognitive-Domain Differences by Cycle and Their Cross-Cycle Change

Contrast	2019 diff (d_{av}/d_z)	2023 diff (d_{av}/d_z)	Δ gap	z	p
Reasoning - Applying	+12.5 (0.117/0.421)***	+1.6 (0.015/0.056)	-10.9	-5.54	< .001
Reasoning - Knowing	+9.7 (0.082/0.255)***	+6.7 (0.058/0.186)***	-3.0	-1.31	.189
Applying - Knowing	-2.8 (-0.025/-0.072)*	+5.1 (0.044/0.157)***	+7.9	4.61	< .001

Note. Positive values indicate the first domain exceeds the second. d = paired Cohen's d . * $p < .05$. ** $p < .01$. *** $p < .001$.

Predictors of Cognitive-Domain Achievement

The three cognitive domains were strongly intercorrelated ($r_s = .95-.97$; Table 2). Among the context measures, mathematics self-confidence showed the strongest bivariate associations with achievement ($r_s = .44-.50$), followed by home educational resources ($r_s = .43-.48$) and liking mathematics ($r_s = .25-.31$), whereas gender was essentially unrelated to achievement at the bivariate level ($|r_s| \leq .07$). Self-confidence and liking mathematics were themselves highly correlated ($r = .70$ in 2019, $.67$ in 2023), which anticipates the suppression pattern observed in the multivariable models.

Table 5 and Figure 2 present the standardized regression coefficients. Across all three domains and both cycles, mathematics self-confidence was the strongest predictor ($\beta = .42-.45$, $p < .001$), followed by home educational resources ($\beta = .33-.41$, $p < .001$). Once self-confidence was held constant, liking mathematics carried a negligible or slightly negative unique association with achievement (β from $-.05$ to $.00$), consistent with its high overlap with self-confidence. Girls scored marginally higher than boys, an association that was clearest for knowing and reasoning in 2019 ($\beta = .11-.13$) and attenuated toward zero by 2023. The models accounted for 35–40% of the variance in domain achievement ($R^2 = .35-.40$). Notably, the predictor structure was almost identical across the three domains: no domain was distinguished by a different leading predictor, and self-confidence and home resources dominated knowing, applying, and reasoning alike.

Table 5

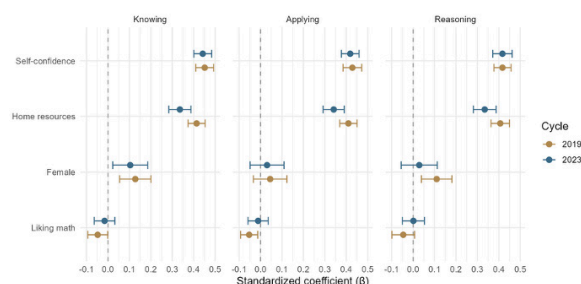
Standardized Regression Coefficients Predicting Cognitive-Domain Achievement by Cycle

Domain	Predictor	β 2019	β 2023	$\Delta\beta$	p ($\Delta\beta$)
Knowing	Self-confidence	.45***	.44***	-0.01	.771
	Home resources	.41***	.34***	-0.08	.019
	Female	.13***	.10*	-0.02	.670
	Liking mathematics	-.05*	-.02	+0.03	.347
Applying	Self-confidence	.43***	.42***	-0.01	.735
	Home resources	.41***	.34***	-0.07	.034
	Female	.04	.03	-0.01	.799
	Liking mathematics	-.05**	-.01	+0.04	.190
Reasoning	Self-confidence	.42***	.42***	+0.00	.995
	Home resources	.41***	.33***	-0.07	.037
	Female	.11**	.03	-0.08	.149
	Liking mathematics	-.05	.00	+0.05	.211

Note. Standardized coefficients from separate PV-pooled, jackknifed regressions per domain and cycle; all predictors entered simultaneously. $\Delta\beta = \beta_{2023} - \beta_{2019}$ (independent-samples test). Model R^2 (2019/2023): Knowing = 0.40/0.36; Applying = 0.37/0.35; Reasoning = 0.36/0.35. * $p < .05$. ** $p < .01$. *** $p < .001$.

Figure 2

Predictors of Cognitive-Domain Achievement by Domain and Cycle, Türkiye Grade 8



Note. Standardized coefficients (β) with 95% confidence intervals from separate PV-pooled, jackknifed models per domain. Cycles are offset vertically for legibility.

Cross-cycle comparison of the coefficients revealed only one systematic change: the standardized effect of home educational resources declined modestly but significantly across all three domains ($\Delta\beta \approx -.07$ to $-.08$, $ps < .05$), while the coefficients for self-confidence, gender, and liking mathematics were stable (all $ps > .14$). Accordingly, the relative shortfall in knowing that emerged by 2023 cannot be attributed to a domain-specific configuration of student-level predictors; the same affective and resource factors predicted all three cognitive domains to a very similar degree in both cycles.

Multilevel robustness. To account for the nesting of students within schools, we re-estimated the predictor models as two-level random-intercept models. Between-school variation was substantial (ICC = .36 in 2019 and .33 in 2023), yet the central findings were unchanged: mathematics self-confidence remained the strongest predictor of overall achievement ($\beta = .43$ and $.40$), the gender coefficient was stable, and liking mathematics remained negligible. The one coefficient sensitive to the multilevel specification was home educational resources, which attenuated from $\beta = .42$ to $.21$ in 2019 and from $.35$ to $.17$ in 2023. A contextual model clarified this pattern: the within-school (individual) coefficient was small ($\beta = .17$ and $.15$), whereas the school-mean coefficient was large ($\beta = .36$ and $.33$), indicating that the home-resources association operates substantially through between-school composition rather than at the individual level.

Discussion

This study profiled eighth-grade mathematics achievement in Türkiye across the knowing, applying, and reasoning domains in two TIMSS cycles and examined how affective and resource factors relate to each domain. Four expectations guided the analysis. Consistent with H1, reasoning differed from the more routine domains within each cycle, though the differences were small; consistent with H2, the cognitive-domain profile changed between 2019 and 2023 rather than rising uniformly; consistent with H3, mathematics self-confidence and home educational resources were the dominant correlates in every domain while liking mathematics added little unique variance; and consistent with H4, the configuration of predictors was essentially invariant across the three domains. We develop these findings in turn, giving particular attention to what they can and cannot support given the observational design and the measurement properties of cognitive subscores.

No reasoning deficit. Contrary to a widespread assumption, reasoning was not the weakest domain in either cycle. In 2019 it was the relative strength, exceeding applying by 12.5 score points ($d_{av} = 0.12$) and knowing by 9.7 points ($d_{av} = 0.08$); by 2023 the three domains had converged, with reasoning

and applying statistically indistinguishable and knowing the relative low point. Two features of this result warrant emphasis. First, although several within-cycle differences were statistically reliable, all were small in effect-size terms ($|d| \leq 0.12$), so the domains are better described as close in standing than as sharply differentiated—a reading reinforced by their very high intercorrelations ($r_s = .95-.97$). Statistical significance here reflects the large sample and the strong within-student correlation among domains as much as any substantive gap. Second, the ordering in which reasoning is not a weakness is not new for Türkiye. The Ministry's own TIMSS 2019 reporting noted comparatively strong reasoning performance (MoNE, 2020), and the present estimates place that observation on an inferential footing. The pattern qualifies, without wholly overturning, the familiar “strong on procedures, weak on reasoning” characterization: analyses of the intended Turkish middle-school curriculum indicate that reasoning demands are nominally well represented, particularly at Grade 8 (Bulut & Taşpınar Şener, 2023), whereas studies of textbooks and classroom tasks document a predominance of low-demand, procedure-oriented work (Polat & Dede, 2023; Horzum & Duran, 2024). The proceduralist narrative may therefore describe the implemented curriculum more accurately than the intended one. Two caveats temper the reframing. The convergence observed here is not unique to Türkiye. Internationally, reasoning shifted from a common relative strength in 2019 to a common relative weakness in 2023 (Mullis et al., 2020; von Davier et al., 2024), so Türkiye's trajectory partly tracks a global movement. And because TIMSS cognitive subscales rest on item subsets and are less robust than the overall score, the profile—consistent with the domain-profiling approach adopted elsewhere (Mokgwathi, 2026)—should be read as indicative rather than definitive.

A shifting profile, led by applying. The 2019–2023 improvement was concentrated in the applying domain, which rose 18.4 points ($d = 0.17$) and was the only domain whose cross-cycle gain reached significance; the gains in knowing (+10.6) and reasoning (+7.5) were positive but not reliable. Correspondingly, applying moved from slightly below to significantly above knowing, and the reasoning advantage over applying narrowed. The profile thus changed in shape rather than simply shifting upward. This redistribution warrants one clarification: because knowing achievement still rose numerically (+10.6 points), its emergence as the relative low point reflects a slower rate of improvement, not a decline in factual or procedural competence—reading it as deterioration would overstate what the data support. Three readings compete to explain the change. The most plausible is the cumulative influence of the skill-based assessment environment that expanded

following the introduction of Türkiye's high-stakes transition examination, in which mathematics items increasingly emphasized contextual interpretation, inference, and the functional use of knowledge over routine recall (Ekinçi & Bal, 2019; Erden, 2020; Uzun & Ağaç, 2023; Bozkuş & Özgeldi, 2024). This is better characterized as a sustained assessment and instructional emphasis than as a formal curricular reform, since Türkiye's most recent curriculum revision postdates the 2023 administration. A pandemic reading is more equivocal. The 2023 cohort's schooling was disrupted by closures associated with reduced participation and mathematics learning loss in Türkiye (Haser et al., 2022) and with substantial but heterogeneous losses internationally (Gajderowicz et al., 2026), so the applying gains might signal selective recovery—but the relative weakness in knowing cannot be attributed specifically to the pandemic with these data. An instrumentation reading is weaker than it first appears: the Grade 8 cognitive-domain target weights were unchanged across cycles (Mullis & Martin, 2017), and because Türkiye had already administered the digital assessment in 2019 (Yalçın, 2022), a simple paper-to-digital mode effect is unlikely (Fishbein et al., 2018). Even so, TIMSS 2023 strengthened its use of contextualized, interactive items and a group-adaptive design (von Davier et al., 2024), so differences in item context and composition remain a residual consideration alongside the subscale-reliability caveat.

Self-confidence and home resources dominate; liking adds little. Across every domain and both cycles, mathematics self-confidence was the strongest correlate of achievement ($\beta = .42-.45$), followed by home educational resources ($\beta = .33-.41$); once self-confidence was controlled, liking mathematics carried a negligible or slightly negative unique association (β from $-.05$ to $.00$). Because self-confidence and liking were themselves highly correlated ($r = .70$ in 2019, $.67$ in 2023), this pattern is consistent with suppression rather than with enjoyment being irrelevant to achievement. The dominance of self-confidence and home resources aligns with prior Turkish TIMSS work (Ersan & Rodriguez, 2020; Coşkun & Karakaya Özyer, 2023) and with meta-analytic evidence for a robust confidence–achievement association (Çiftçi & Yıldız, 2019). Interpreting it, however, requires care about what TIMSS measures. The confidence scale indexes a competence-related self-belief closer to academic self-concept—normative and comparatively stable—than to task-specific self-efficacy (Bong & Skaalvik, 2003), and self-concept and achievement are reciprocally related over time (Marsh & Craven, 2006; Marsh & Martin, 2011). Cross-sectional dominance therefore cannot establish causal direction. Nor should the weak adjusted coefficient for liking be read as showing enjoyment to be educationally inert: situated expectancy–value theory assigns competence beliefs

a more proximal role in performance while enjoyment remains consequential for engagement, persistence, and later choices (Eccles & Wigfield, 2020), and the attenuation of a liking effect once confidence is entered is a recurring TIMSS pattern (Østbø & Zachrisson, 2022; Ivanova & Michaelides, 2022). On balance, the evidence points—tentatively—toward prioritizing competence-related confidence through mastery experiences and strategy-focused feedback while continuing to cultivate interest (Usher & Pajares, 2009). Because the design cannot show confidence to be causally prior, this is best read as a direction for intervention research rather than a warranted prescription.

The same factors predict all domains. The predictor configuration was essentially invariant across knowing, applying, and reasoning: no domain was distinguished by a different leading correlate, and the models accounted for similar variance throughout ($R^2 = .35-.40$). A comparably domain-general pattern has been reported for Turkish fourth-graders in TIMSS 2019 (Ayva Yörü et al., 2023). This invariance is informative, but not in the way it may first appear: a predictor set that does not vary across domains cannot, by construction, account for variation between them. Student-level affect and resources are therefore silent on the shortfall in knowing that emerged by 2023—not because they were tested and failed, but because they are the wrong grain of analysis for a domain-specific gap. That question belongs more properly to students' opportunities to learn. At the level of the intended curriculum, knowing appears under-weighted relative to the TIMSS Grade 8 targets, while reasoning is nominally well represented but concentrated in lower-order analytic outcomes, with generalizing and justifying scarcely present (Bulut & Taşpınar Şener, 2023); central examinations have moved in a similar direction, with knowledge-level items declining across cycles (Bostan-Sarioğlu et al., 2021; Ekinci & Bal, 2019; Uzun & Ağa, 2023). A relative weakness in knowing is thus less an anomaly to be explained away than a plausible consequence of what students are given the opportunity to learn (Schmidt et al., 1997). Evidence on the enacted curriculum is less consistent: textbook and classroom studies document substantial reliance on low-demand procedural tasks and teacher-led instruction (Şensoy & Kılıç, 2021; Polat & Dede, 2023; Horzum & Duran, 2024), teachers report gaps between skill-based assessment demands and available curricular guidance (Erden, 2020; Kertil et al., 2021), and teacher quality and working conditions account for meaningful between-school variation in achievement (Sezer & Çakan, 2022). This account remains provisional; TIMSS itself cautions that content coverage is only one component of opportunity to learn and that enacted curricula frequently diverge from intended ones (Bakan Kalaycioglu et al., 2026). Establishing it would require linking domain-level

performance to teacher-reported coverage rather than inferring it from curriculum documents alone.

Small, diminishing gender gaps. Gender differences were small and, where detectable, favored girls in knowing and reasoning in 2019 ($\beta = .11-.13$), attenuating toward zero by 2023. This limited differentiation is consistent with earlier Turkish analyses in which gender did not significantly predict overall mathematics achievement net of student and school characteristics (Coşkun & Karadağ, 2023; Şahin et al., 2025) and with the international finding of broad parity at Grade 8 (Mullis et al., 2020); cross-national work likewise indicates that mathematics gender gaps are small and variable in magnitude and direction (Else-Quest et al., 2010; Meinck & Brese, 2019). The attenuation admits a substantive reading rather than mere regression to the mean: pandemic-related learning losses were not gender-neutral, falling disproportionately on girls in international analyses (Gajderowicz et al., 2026), so the narrowing may reflect an erosion of girls' prior advantage rather than gains among boys—a distinction these data cannot adjudicate but which future cycles should. Achievement parity should not, in any case, be equated with affective parity: TIMSS-based studies find that girls often report less positive mathematics self-beliefs even where their achievement matches or exceeds that of boys (Ghasemi & Burley, 2019; Mejía-Rodríguez et al., 2021), a gap with potential implications for later participation in advanced mathematics and STEM.

Achievement inequality is largely a between-school phenomenon. The multilevel models added an important qualification. Although home educational resources predicted achievement strongly in single-level models, most of this association operated between schools rather than within them (within-school $\beta \approx .15-.17$ vs. school-mean $\beta \approx .33-.36$), against a backdrop of substantial between-school variance ($ICC \approx .33-.36$). Students' achievement was thus bound up not only with their own resources but with the resource composition of the school they attended. This is consistent with earlier Turkish TIMSS evidence that socioeconomic status predicts achievement more strongly at the school than the student level (Ersan & Rodriguez, 2020) and that a substantial share of eighth-grade achievement variance lies between schools (Suna & Özer, 2021), as well as with national and international findings that students of differing backgrounds are unevenly distributed across school types closely tied to composition, resources, and learning environments (Alacacı & Erbaş, 2010; Özdemir, 2016; Suna et al., 2020; Özer & Suna, 2023). The mechanism, however, requires care: because these are Grade 8 students who have not yet entered upper-secondary tracks, the finding cannot be read as evidence that high-school tracking produced the observed variance. Rather, it situates achievement

within a system in which between-school differences emerge before upper-secondary placement and are subsequently reinforced by selective transitions (Bölükbaş & Gür, 2020; Niehues et al., 2019). The policy implication is that narrowing achievement gaps is unlikely to be achieved through student-level interventions alone and may require attention to school composition and the distribution of resources across schools. Mathematics self-confidence, by contrast, remained a robust individual-level correlate and is tempting to read as a malleable within-school lever; that reading warrants caution, since the confidence scale indexes academic self-concept, which is itself compositionally shaped—equally able students tending to report lower self-concept in higher-achieving settings (Marsh & Craven, 2006; Bong & Skaalvik, 2003).

Limitations and future directions

Several constraints qualify these findings. The design is observational and cross-sectional within each cycle; although single-level estimates are complemented by multilevel models, no causal claims follow, and this matters most for the plausibly bidirectional confidence–achievement association (Marsh & Craven, 2006). Measurement invariance of the context scales was assumed rather than formally tested across cycles and gender, so apparent change in an affective predictor could partly reflect a shift in measurement rather than in the construct (Atılğan & Deniz, 2023). The cycles also differ instrumentally: TIMSS 2023 completed the transition to a fully digital, group-adaptive assessment with new interactive formats (von Davier et al., 2024); although Türkiye's prior digital administration makes a simple mode effect unlikely (Yalçın, 2022), differences in item context remain a possible confound for profile comparison. The profile rests on two cycles only, so the convergence reported here cannot yet be distinguished from cycle-specific fluctuation; and the 2023 cohort's schooling was disrupted unevenly by the pandemic (Gajderowicz et al., 2026), so cohort effects are not fully separable from trend. Finally, cognitive-domain subscores rest on smaller item subsets and are measured less precisely than the overall score, so the small differences among knowing, applying, and reasoning should be interpreted alongside the reported effect sizes and confidence intervals rather than as substantive findings in their own right. These constraints converge on a common remedy: longitudinal or cross-lagged designs to establish temporal ordering, formal invariance testing before cross-cycle and cross-group comparison, and replication on mode-invariant trend items across additional cycles.

Conclusion

Disaggregating Turkish eighth-graders' mathematics achievement by cognitive domain yields a picture

that overall scores obscure: reasoning is not a national weakness, the recent improvement was led by applying while knowing lagged, and the same affective and resource factors predict all three domains alike. Because those student-level factors cannot, by construction, explain a domain-specific gap, the emerging shortfall in knowing points instead toward opportunity-to-learn, and the strong between-school patterning of resources locates much of the inequality at the level of schools rather than individuals.

More broadly, these findings illustrate the value of examining mathematics achievement through its cognitive dimensions rather than overall scores alone. Overall achievement remains an indispensable indicator for monitoring educational systems, but a domain-level lens reveals patterns—an applying-led rise, a knowing shortfall unexplained by student-level factors, and inequality concentrated between schools—those composite scores leave hidden. Read with due caution about the observational design and the reliability of subscores, the results suggest that both curricular alignment and the equitable distribution of school resources deserve closer attention than overall achievement trends alone would imply, and that pairing cognitive-domain achievement with classroom, teacher, and curriculum data is a priority for future research.

References

- Akyüz, G. (2014). The effects of student and school factors on mathematics achievement in TIMSS 2011. *Education and Science*, 39(172), 150–162. <https://doi.org/10.15390/ES.2014.1255>
- Alacacı, C., & Erbaş, A. K. (2010). Unpacking the inequality among Turkish schools: Findings from PISA 2006. *International Journal of Educational Development*, 30(2), 182–192. <https://doi.org/10.1016/j.ijedudev.2009.03.006>
- Anderson, L. W., & Krathwohl, D. R. (Eds.). (2001). *A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives* (Complete ed.). Longman.
- Arıkan, S., Özer, F., Şeker, V., & Ertaş, G. (2020). The importance of sample weights and plausible values in large-scale assessments. *Journal of Measurement and Evaluation in Education and Psychology*, 11(1), 43–60. <https://doi.org/10.21031/epod.602765>
- Atılğan, M., & Deniz, K. Z. (2023). Investigation of the measurement invariance of affective characteristics related to TIMSS 2019 mathematics achievement by gender. *Journal of Measurement and Evaluation in Education and Psychology*, 14(3), 185–199. <https://doi.org/10.21031/epod.1221365>

- Ayva Yörü, F. G., Sezer Başaran, E., & Çakan, M. (2023). Okul, ev ve öğrenci özelliklerinin 4. sınıf öğrencilerinin TIMSS 2019 matematik bilişsel becerilerindeki rolü [The role of school, home and student characteristics in TIMSS 2019 mathematics cognitive skills of 4th grade students]. *Mehmet Akif Ersoy University Journal of Education Faculty*, 67, 197–227. <https://doi.org/10.21764/mauefd.1142768>
- Bakan Kalaycioglu, D., Reynolds, K. A., Fishbein, B., Gonzalez, E., & von Davier, M. (2026). *Intended, implemented, and attained mathematics and science curricula at the eighth grade: Findings from TIMSS 2023*. TIMSS & PIRLS International Study Center, Boston College. <https://doi.org/10.6017/lse.tpisc.timss.wf9486>
- Ball, D. L., & Bass, H. (2003). Making mathematics reasonable in school. In J. Kilpatrick, W. G. Martin, & D. Schifter (Eds.), *A research companion to Principles and Standards for School Mathematics* (pp. 27–44). National Council of Teachers of Mathematics.
- Bölükbaş, S., & Gür, B. S. (2020). Tracking and inequality: The results from Turkey. *International Journal of Educational Development*, 78, Article 102262. <https://doi.org/10.1016/j.ijedudev.2020.102262>
- Bong, M., & Skaalvik, E. M. (2003). Academic self-concept and self-efficacy: How different are they really? *Educational Psychology Review*, 15(1), 1–40. <https://doi.org/10.1023/A:1021302408382>
- Bostan-Sarioğlu, A., Dolu, G., & Sevim, N. (2021). Sekizinci sınıf merkezi sınavlardaki fen sorularının TIMSS-2019 bilişsel alanlarına göre analizi [Analysis of science questions in eighth grade central exams according to cognitive fields of TIMSS-2019]. *e-Kafkas Journal of Educational Research*, 8(3), 514–533. <https://doi.org/10.30900/kafkasegt.973021>
- Bozkuş, Ç., & Özgeldi, M. (2024). Analysis of middle school mathematics teachers' evaluations of skill-based questions within the framework of mathematical competencies. *International Journal of Turkish Education Sciences*, 12(1), 535–583. <https://doi.org/10.46778/goputeb.1425323>
- Braun, H., & von Davier, M. (2017). The use of test scores from large-scale assessment surveys: Psychometric and statistical considerations. *Large-scale Assessments in Education*, 5(1), Article 17. <https://doi.org/10.1186/s40536-017-0050-x>
- Bulut, A. S., & Taşpınar Şener, Z. (2023). Analysis of secondary school mathematics curriculum learning outcomes by TIMSS-2019 cognitive domain skills. *Türk Akademik Yayınlar Dergisi (TAY Journal)*, 7(1), 303–328. <https://doi.org/10.29329/tayjournal.2023.53714>
- Choi, K. M., Mun, J., & Hand, B. (2026). Bridging reasoning in mathematics and science: A document analysis of applying and reasoning in the TIMSS assessment framework and the case for utilization of resources. *Educational Psychology Review*, 38, Article 22. <https://doi.org/10.1007/s10648-025-10115-2>
- Çiftçi, Ş. K., & Yıldız, P. (2019). The effect of self-confidence on mathematics achievement: The meta-analysis of Trends in International Mathematics and Science Study (TIMSS). *International Journal of Instruction*, 12(2), 683–694. <https://doi.org/10.29333/iji.2019.12243a>
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Lawrence Erlbaum Associates.
- Coşkun, B., & Karadağ, E. (2023). The effect of student and school characteristics on TIMSS 2015 science and mathematics achievement: The case of Türkiye. *Journal of Pedagogical Research*, 7(1), 203–227. <https://doi.org/10.33902/JPR.202318875>
- Coşkun, B., & Karakaya Özyer, K. (2023). The effect of student characteristics and socioeconomic status on mathematics achievement in Türkiye: Insights from TIMSS 2011-2019. *International Journal of Assessment Tools in Education*, 10(3), 454–481. <https://doi.org/10.21449/ijate.1272517>
- Eccles, J. S., & Wigfield, A. (2020). From expectancy-value theory to situated expectancy-value theory: A developmental, social cognitive, and sociocultural perspective on motivation. *Contemporary Educational Psychology*, 61, Article 101859. <https://doi.org/10.1016/j.cedpsych.2020.101859>
- Ekinci, O., & Bal, A. P. (2019). 2018 yılı Liseye Geçiş Sınavı (LGS) matematik sorularının öğrenme alanları ve Yenilenmiş Bloom Taksonomisi bağlamında değerlendirilmesi [Evaluation of High School Entrance Exam (LGS) 2018 in terms of mathematics learning field and Revised Bloom Taxonomy]. *Anemon Muş Alparslan Üniversitesi Sosyal Bilimler Dergisi*, 7(3), 9–18. <https://doi.org/10.18506/anemon.462717>

- Else-Quest, N. M., Hyde, J. S., & Linn, M. C. (2010). Cross-national patterns of gender differences in mathematics: A meta-analysis. *Psychological Bulletin*, 136(1), 103–127. <https://doi.org/10.1037/a0018053>
- Erden, B. (2020). Türkçe, matematik ve fen bilimleri dersi beceri temelli sorularına ilişkin öğretmen görüşleri [Teachers' views related to skill-based questions in Turkish, mathematics and science lessons]. *Academia Journal of Educational Research*, 5(2), 270–292. <https://izlik.org/JA67NL34CY>
- Ersan, Ö., & Rodriguez, M. C. (2020). Socioeconomic status and beyond: A multilevel analysis of TIMSS mathematics achievement given student and school context in Turkey. *Large-scale Assessments in Education*, 8, Article 15. <https://doi.org/10.1186/s40536-020-00093-y>
- Fishbein, B., Martin, M. O., Mullis, I. V. S., & Foy, P. (2018). The TIMSS 2019 item equivalence study: Examining mode effects for computer-based assessment and implications for measuring trends. *Large-scale Assessments in Education*, 6, Article 11. <https://doi.org/10.1186/s40536-018-0064-z>
- Foy, P., & LaRoche, S. (2020). Estimating standard errors in the TIMSS 2019 results. In M. O. Martin, M. von Davier, & I. V. S. Mullis (Eds.), *Methods and procedures: TIMSS 2019 technical report*. TIMSS & PIRLS International Study Center, Boston College
- Gajderowicz, T., Jakubowski, M., Kennedy, A. I., Kjeldsen, C. C., Patrinos, H. A., & Strietholt, R. (2026). The learning crisis: Three years after COVID-19. *Learning and Instruction*, 104, Article 102373. <https://doi.org/10.1016/j.learninstruc.2026.102373>
- Ghasemi, E., & Burley, H. (2019). Gender, affect, and math: A cross-national meta-analysis of Trends in International Mathematics and Science Study 2015 outcomes. *Large-scale Assessments in Education*, 7, Article 10. <https://doi.org/10.1186/s40536-019-0078-1>
- Haser, Ç., Doğan, O., & Kurt Erhan, G. (2022). Tracing students' mathematics learning loss during school closures in teachers' self-reported practices. *International Journal of Educational Development*, 88, Article 102536. <https://doi.org/10.1016/j.ijedudev.2021.102536>
- Hernández-Torrano, D., & Courtney, M. G. R. (2021). Modern international large-scale assessment in education: An integrative review and mapping of the literature. *Large-scale Assessments in Education*, 9, Article 17. <https://doi.org/10.1186/s40536-021-00109-1>
- Horzum, T., & Duran, A. A. (2024). An examination of the cognitive demand levels reflected by mathematics tasks in the middle grades mathematics textbooks. *International Journal of Educational Spectrum*, 6(2), 294–321. <https://doi.org/10.47806/ijesacademic.1522050>
- Ivanova, M., & Michaelides, M. P. (2022). Motivational components in TIMSS 2015 and their effects on engaging teaching practices and mathematics performance. *Studies in Educational Evaluation*, 74, Article 101173. <https://doi.org/10.1016/j.stueduc.2022.101173>
- Kertil, M., Gülbağcı Dede, H., & Ulusoy, E. G. (2021). Skill-based mathematics questions: What do middle school mathematics teachers think about and how do they implement them? *Turkish Journal of Computer and Mathematics Education*, 12(1), 151–186. <https://doi.org/10.16949/turkbilmat.774651>
- Kilpatrick, J., Swafford, J., & Findell, B. (Eds.). (2001). *Adding it up: Helping children learn mathematics*. National Academy Press.
- Koyuncu, İ. (2021). TIMSS international benchmarks of eighth graders in mathematics: A correspondence analysis study. *International Electronic Journal of Elementary Education*, 14(2), 179–194. <https://doi.org/10.26822/iejee.2022.237>
- Lee, J., & Stankov, L. (2018). Non-cognitive predictors of academic achievement: Evidence from TIMSS and PISA. *Learning and Individual Differences*, 65, 50–64. <https://doi.org/10.1016/j.lindif.2018.05.009>
- Marsh, H. W., & Craven, R. G. (2006). Reciprocal effects of self-concept and performance from a multidimensional perspective: Beyond seductive pleasure and unidimensional perspectives. *Perspectives on Psychological Science*, 1(2), 133–163. <https://doi.org/10.1111/j.1745-6916.2006.00010.x>
- Marsh, H. W., & Martin, A. J. (2011). Academic self-concept and academic achievement: Relations and causal ordering. *British Journal of Educational Psychology*, 81(1), 59–77. <https://doi.org/10.1348/000709910X503501>
- Meinck, S., & Brese, F. (2019). Trends in gender gaps: Using 20 years of evidence from TIMSS. *Large-scale Assessments in Education*, 7, Article 8. <https://doi.org/10.1186/s40536-019-0076-3>

- Mejía-Rodríguez, A. M., Luyten, H., & Meelissen, M. R. M. (2021). Gender differences in mathematics self-concept across the world: An exploration of student and parent data of TIMSS 2015. *International Journal of Science and Mathematics Education*, 19(6), 1229–1250. <https://doi.org/10.1007/s10763-020-10100-x>
- Ministry of National Education. (2020). *TIMSS 2019 Türkiye preliminary report* [TIMSS 2019 Türkiye ön raporu]. Millî Eğitim Bakanlığı. https://www.meb.gov.tr/meb_iys_dosyalar/2020_12/10173505_No15_TIMSS_2019_Turkiye_On_Raporu_Guncel.pdf
- Ministry of National Education. (2024). *TIMSS 2023 Türkiye report* [TIMSS 2023 Türkiye raporu]. Directorate General for Measurement, Evaluation and Examination Services. <https://odsgm.meb.gov.tr/www/raporlar/kategori/96>
- Mokgwathi, M. S. (2026). Closing the gap in early mathematics: Domain and cognitive insights from TIMSS 2023 in South Africa and Singapore. *F1000Research*, 14, Article 1209. <https://doi.org/10.12688/f1000research.172015.3>
- Mullis, I. V. S., & Martin, M. O. (Eds.). (2017). *TIMSS 2019 assessment frameworks*. TIMSS & PIRLS International Study Center, Boston College.
- Mullis, I. V. S., Martin, M. O., & von Davier, M. (Eds.). (2021). *TIMSS 2023 assessment frameworks*. TIMSS & PIRLS International Study Center, Boston College.
- Mullis, I. V. S., Martin, M. O., Foy, P., Kelly, D. L., & Fishbein, B. (2020). *TIMSS 2019 international results in mathematics and science*. TIMSS & PIRLS International Study Center, Boston College.
- Niehues, W., Kisbu-Sakarya, Y., & Selçuk, B. (2019). Implications of between-school tracking for Turkish students. *Turkish Journal of Education*, 8(3), 196–216. <https://doi.org/10.19128/turje.453383>
- Niss, M. (2003). Mathematical competencies and the learning of mathematics: The Danish KOM project. In A. Gagatsis & S. Papastavridis (Eds.), *Proceedings of the 3rd Mediterranean Conference on Mathematical Education* (pp. 115–124). Hellenic Mathematical Society.
- Oberleiter, S., Fries, J., Schock, L. S., Steininger, B., & Pietschnig, J. (2023). Predicting cross-national sex differences in large-scale assessments of students' reading literacy, mathematics, and science achievement: Evidence from PIRLS and TIMSS. *Intelligence*, 100, Article 101784. <https://doi.org/10.1016/j.intell.2023.101784>
- Østbø, I. U., & Zachrisson, H. D. (2022). Student motivation and parental attitude as mediators for SES effects on mathematics achievement: Evidence from Norway in TIMSS 2015. *Scandinavian Journal of Educational Research*, 66(5), 808–823. <https://doi.org/10.1080/00313831.2021.1939138>
- Özdemir, C. (2016). Equity in the Turkish education system: A multilevel analysis of social background influences on the mathematics performance of 15-year-old students. *European Educational Research Journal*, 15(2), 193–217. <https://doi.org/10.1177/1474904115627159>
- Özer, M., & Suna, E. (2023). The relationship between school socioeconomic composition and academic achievement in Türkiye. *Journal of Economy Culture and Society*, 66, 17–27. <https://doi.org/10.26650/JECS2022-1116979>
- Philpot, R., Lindquist, M., Mullis, I. V. S., & Aldrich, C. E. A. (2021). TIMSS 2023 mathematics framework. In I. V. S. Mullis & M. O. Martin (Eds.), *TIMSS 2023 assessment frameworks*. (pp. 5-19). Boston College, TIMSS & PIRLS International Study Center <https://timssandpirls.bc.edu/timss2023/frameworks/index.html>
- Polat, S., & Dede, Y. (2023). Trends in cognitive demands levels of mathematical tasks in Turkish middle school mathematics textbooks: Algebra learning domain. *International Journal for Mathematics Teaching and Learning*, 24(1), 40–61. <https://doi.org/10.4256/ijmtl.v24i1.476>
- R Core Team. (2025). *R: A language and environment for statistical computing* [Computer software]. R Foundation for Statistical Computing. <https://www.R-project.org/>
- Reilly, D., Neumann, D. L., & Andrews, G. (2019). Investigating gender differences in mathematics and science: Results from the 2011 Trends in Mathematics and Science Survey. *Research in Science Education*, 49, 25–50. <https://doi.org/10.1007/s11165-017-9630-6>
- Rubin, D. B. (1987). *Multiple imputation for nonresponse in surveys*. Wiley. <https://doi.org/10.1002/9780470316696>
- Şahin, A., Kurt, G., & Yerdelen-Damar, S. (2025). Factors predicting Turkish students' mathematics and science achievement in TIMSS 2019. *Electronic Journal for Research in Science & Mathematics Education*, 29(1), 42–59.

- Schmidt, W. H., McKnight, C. C., Valverde, G. A., Houang, R. T., & Wiley, D. E. (Eds.). (1997). *Many visions, many aims: A cross-national investigation of curricular intentions in school mathematics* (Vol. 1). Kluwer Academic Publishers. <https://doi.org/10.1007/978-94-011-5786-5>
- Şensoy, Ç. P., & Kılıç, A. (2021). Ortaokul matematik öğretmenlerinin ders işleyiş süreçlerinin incelenmesi [Analysis of the middle school math teachers' teaching processes]. *Bolu Abant İzzet Baysal Üniversitesi Eğitim Fakültesi Dergisi*, 21(2), 434–452. <https://doi.org/10.17240/aibuefd.2021.21.62826-874046>
- Sezer, E., & Çakan, M. (2022). Role of teacher quality and working conditions in TIMSS 2019 mathematics achievement. *Journal of Theoretical Educational Sciences*, 15(2), 395–419. <https://doi.org/10.30831/akukeg.971286>
- Suna, H. E., & Özer, M. (2021). Türkiye'de sosyoekonomik düzey ve okullar arası başarı farklarının akademik başarı ile ilişkisi [The achievement gap between schools and relationship between achievement and socioeconomic status in Turkey]. *Journal of Measurement and Evaluation in Education and Psychology*, 12(1), 54–70. <https://doi.org/10.21031/epod.860431>
- Suna, H. E., Tanberkan, H., Gür, B. S., Perc, M., & Özer, M. (2020). Socioeconomic status and school type as predictors of academic achievement. *Journal of Economy Culture and Society*, 61, 41–64. <https://doi.org/10.26650/JECS2020-0034>
- Usher, E. L., & Pajares, F. (2009). Sources of self-efficacy in mathematics: A validation study. *Contemporary Educational Psychology*, 34(1), 89–101. <https://doi.org/10.1016/j.cedpsych.2008.09.002>
- Uzun, H., & Ağaç, G. (2023). Ortaokul matematik öğretmenlerinin beceri temelli sorulara ilişkin yaklaşımlarının incelenmesi [An investigation of secondary school mathematics teachers' approaches to skill-based question]. *İnönü University Journal of the Faculty of Education*, 24(1), 57–78. <https://doi.org/10.17679/inuefd.1133896>
- von Davier, M., Kennedy, A., Reynolds, K., Fishbein, B., Khorramdel, L., Aldrich, C., Bookbinder, A., Bezirhan, U., & Yin, L. (2024). *TIMSS 2023 international results in mathematics and science*. TIMSS & PIRLS International Study Center, Boston College. <https://doi.org/10.6017/lse.tpisc.timss.rs6460>
- Yalçın, S. (2022). Examining students' item response times in eTIMSS according to their proficiency levels, self-confidence, and item characteristics. *Journal of Measurement and Evaluation in Education and Psychology*, 13(1), 23–39. <https://doi.org/10.21031/epod.999545>
- Yıldırım, S. (2019). Predicting mathematics achievement: The role of socioeconomic status, parental involvement, and self-confidence. *Education and Science*, 44(198), 99–113. <https://doi.org/10.15390/EB.2019.7868>
- Yin, L., & Foy, P. (2021). TIMSS 2023 assessment design. In I. V. S. Mullis & M. O. Martin (Eds.), *TIMSS 2023 assessment frameworks*. (pp. 71–85). Boston College, TIMSS & PIRLS International Study Center <https://timssandpirls.bc.edu/timss2023/frameworks/index.html>
- Zhu, Z. (2023). International comparative study of learning trajectories based on TIMSS 2019 G4 data on cognitive diagnostic models. *Frontiers in Psychology*, 14, 1241656. <https://doi.org/10.3389/fpsyg.2023.1241656>



This page is intentionally left blank.
www.iejee.com

Enhancing Elementary Students' Literacy Engagement through a Web-Based Digital Literature Platform: A Mixed-Methods Study and Instructional Framework

Nurjanah^{a*}, Yurdayanti^b, Yorenza Meifinda^c, Ilpandari^d, Winda Purnama Sari^e

Received : 21 January 2026
 Revised : 16 June 2026
 Accepted : 15 September 2026
 DOI : 10.26822/iejee.2026.468

^{a*} **Corresponding Author:** Nurjanah, Department of Primary Teacher Education, Universitas Muhammadiyah Bangka Belitung, Bangka Belitung, Indonesia
 E-mail: nurjanah@unmuhbabel.ac.id
 ORCID: <https://orcid.org/0009-0000-0604-5937>

^b Yurdayanti, Department of Primary Teacher Education, Universitas Muhammadiyah Bangka Belitung, Bangka Belitung, Indonesia
 E-mail: yurdayanti@unmuhbabel.ac.id
 ORCID: <https://orcid.org/0009-0008-3036-8881>

^c Yorenza Meifinda, Department of Primary Teacher Education, Universitas Muhammadiyah Bangka Belitung, Bangka Belitung, Indonesia
 E-mail: yorenza.meifinda@unmuhbabel.ac.id
 ORCID: <https://orcid.org/0009-0004-3018-4987>

^d Ilpandari, Faculty of Engineering and Science, Universitas Muhammadiyah Bangka Belitung, Bangka Belitung, Indonesia
 E-mail: ilpandari@unmuhbabel.ac.id
 ORCID: <https://orcid.org/0009-0000-4804-9557>

^e Winda Purnama Sari, Department of Educational Science, Graduate School, Universitas Negeri Yogyakarta, Yogyakarta Indonesia
 E-mail: windapurnama.2023@student.uny.ac.id
 ORCID: <https://orcid.org/0000-0001-5785-1023>

Abstract

The increasing integration of digital and multimodal texts underscores the need for more empirical research at the elementary level on how pedagogically designed platforms enhance literacy engagement. To address this gap, this study examined the effectiveness of a web-based digital literature platform in fostering elementary students' literacy engagement through pedagogically mediated instruction. An explanatory sequential mixed-methods design was employed involving 30 elementary students who participated in an eight-week instructional intervention. Quantitative data were collected using pre- and post-tests measuring cognitive, affective, and behavioral engagement, as well as students' abilities to identify key information, analyze intrinsic literary elements, and interpret moral and social meanings in visual and audiovisual texts. Qualitative data from classroom observations, focus group discussions, and teacher interviews were used to elaborate on the quantitative findings. Results showed statistically significant improvements in literacy engagement and multimodal literary literacy. Qualitative evidence indicated that guided multimodal scaffolding, analytical prompts, and reflective activities supported students' progression from a surface-level understanding to an evidence-based interpretation. Based on integrated findings, the study proposes a four-phase instructional framework comprising Engagement, Analysis, Interpretation, and Reflection, offering transferable principles for elementary digital literacy instruction.

Keywords:

digital literature platform; instructional framework; literacy engagement; multimodal literacy; web-based learning



www.iejee.com
 ISSN: 1307-9298

2026 Published by KURA Education & Publishing.
 This is an open access article under the CC BY-NC-ND license. (<https://creativecommons.org/licenses/by/4.0/>)

Introduction

The rapid development of digital technologies has significantly transformed students' literacy practices, shifting from the dominance of print-based texts toward engagement with multimodal texts that integrate visual, audiovisual, and interactive elements (Gnams, 2025; Ylipulli et al., 2023). Within this context, twenty-first-century literacy extends beyond traditional decoding and literal

comprehension. It now encompasses more complex analytical competencies that are essential for students to navigate contemporary texts. These competencies include identifying key information, analyzing intrinsic literary elements (such as theme, plot, and setting), and interpreting moral and social meanings in reflective and argumentative ways (Calafato & Simmonds, 2022; Fradana et al., 2025; Gasser et al., 2025; Walldén & Winlund, 2025).

However, empirical evidence from elementary literature classrooms reveals a substantial gap between curricular expectations and classroom realities (Muzakkir, Saptarina, & Amelia, 2022; Nawi, Syafiq, Nor, & Amani, 2023). Achievement in visual and audiovisual literary literacy remains relatively low, particularly in terms of student engagement and analytical competence (Dahiru, 2020; Najeeb & Jameel, 2024). Classroom observations indicate that many students struggle to identify key information in short audiovisual texts, analyze the intrinsic elements of visual narratives holistically, or formulate logical, evidence-based moral interpretations (Baker et al., 2023; Encabo-Fernández, Albarracín-Vivo, & Jerez-Martínez, 2023). Teachers report that students tend to focus on surface-level visual features—such as color, animation, and character expressions—while structured visual literacy strategies are not yet systematically embedded in instructional planning (Dey & Munshi, 2025; Eryilmaz & Çelik, 2025). At the same time, students often perceive visual narratives primarily as entertainment rather than as objects of critical inquiry, leading to a weak conceptual understanding of literary themes and messages (Gacumo, Coats, & Kucirkova, 2026; Sakanti & Widodo, 2020; Volker Pietsch, 2024).

These difficulties are further evident in students' analytical tasks and classroom discussions. When asked to identify themes, plots, and settings, most students provide fragmented keywords without explaining relationships among elements or supporting their interpretations with multimodal evidence. Classroom discussions rarely reach deeper levels of inquiry and seldom explore multimodal elements such as imagery, background music, point of view, or symbolism. Consequently, instructional goals emphasizing comprehension of literary content, structure, and meaning are not optimally achieved (Najeeb & Jameel, 2024). From an affective and reflective perspective, students tend to recognize only surface-level moral values (e.g., "honesty" or "kindness") without connecting them to broader social and cultural contexts or engaging in ethical reflection through discussion (Muzakkir et al., 2022; Najeeb & Jameel, 2024).

Low achievement in articulating moral, social, and cultural values in literary texts has also been reported in previous studies (Dahiru, 2020). Assessment results

show that students rarely integrate multimodal evidence into value interpretation, while reflective activities and critical discussions remain limited. These conditions are exacerbated by constraints in technological resources, limited teacher training, and the absence of pedagogically grounded multimodal task designs and assessment frameworks (Brosseuk & Downes, 2024; Gil-Flores, Rodríguez-Santero, & Torres-Gordillo, 2017; Hew & Brush, 2007; Nawi et al., 2023).

At the same time, global literacy practices are increasingly oriented toward digital and multimodal forms (Ho, 2024; Kress, 2003). In language and literature education, Digital Multimodal Composing (DMC)—which integrates textual, audio, visual, spatial, and kinetic elements—has gained attention as a pedagogical approach aligned with students' contemporary media cultures (Kessler & Marino, 2023; Sahiruddin, Sudarwati, Lestari, & Daniswara, 2025). Empirical studies indicate that consistent integration of digital multimodal practices can enhance the quality of students' writing in terms of content, organization, and communicative achievement (Maghsoudi, Golshan, & Naeimi, 2022). Nevertheless, empirical and design-based research focusing specifically on elementary students remains limited, particularly studies that test effectiveness using systematic, experimental approaches (Jiang & Hafner, 2025; Mayer, 2024).

These conditions highlight a critical gap between curriculum demands for analytical and reflective literary literacy and classroom practices that still position digital media primarily as playback tools, without structured support for higher-order thinking (Farrar, Arizpe, & Lees, 2024; Mayer, 2024). Such pedagogical insufficiency constrains students' conceptual development, especially in understanding literary themes and constructing moral arguments, thereby impeding the early development of literary reasoning and critical thinking. This issue is particularly salient in visual and audiovisual literacy, given the growing importance of interpreting and evaluating meaning in complex digital media environments as a core twenty-first-century competence (Farrar et al., 2024).

In response to this gap, the present study underscores the urgency of developing pedagogically designed web-based digital literature platforms that facilitate students' cognitive, affective, and behavioral engagement in visual and audiovisual literary learning. Employing a mixed-methods approach, the study examines the platform's effectiveness quantitatively while qualitatively exploring teachers' and students' experiences to achieve a comprehensive understanding of its pedagogical impact.

The novelty of this study lies in: (1) the development of a web-based digital literature platform integrated with

multimodal analytical, interpretative, and reflective tasks; (2) the provision of mixed-methods empirical evidence on the platform's effects on elementary students' literacy engagement; and (3) the formulation of an applicable instructional framework to guide the implementation of digital literary learning in classroom contexts. Accordingly, this study contributes to strengthening evidence-based, sustainable, and pedagogically grounded digital literary practices for improving the quality of elementary literature education in the digital era.

In light of the identified gap between the demands of multimodal literary literacy and current instructional practices in elementary classrooms, this study examines the pedagogical effectiveness of a web-based digital literature platform designed to support structured engagement and higher-order literary meaning-making. Employing an explanatory sequential mixed-methods approach, the study investigates both the outcomes and the instructional processes underlying students' engagement and multimodal literary literacy development. Accordingly, the following research questions guide the study:

RQ 1: Does the use of a pedagogically designed web-based digital literature platform significantly affect elementary students' literacy engagement across cognitive, affective, and behavioral dimensions compared with conventional literature instruction?

RQ 2: Does the use of the platform significantly improve elementary students' multimodal literary literacy in terms of: (a) identifying key information, (b) analyzing intrinsic elements of visual and audiovisual literary texts, and (c) interpreting moral and social meanings in argumentative and reflective ways?

RQ 3: How do students' and teachers' experiences and perceptions explain changes in literacy engagement and analytical literary competence following the use of the web-based digital literature platform?

RQ 4: How do the students utilize multimodal features and tasks within the platform to construct analytical, interpretative, and reflective understandings of visual and audiovisual literary texts?

RQ 5: How can quantitative results and qualitative findings be integrated to formulate an applicable web-based multimodal digital literature instructional framework for elementary education?

Literature Review

Multimodal Digital Literacy and the Challenge of Student Engagement

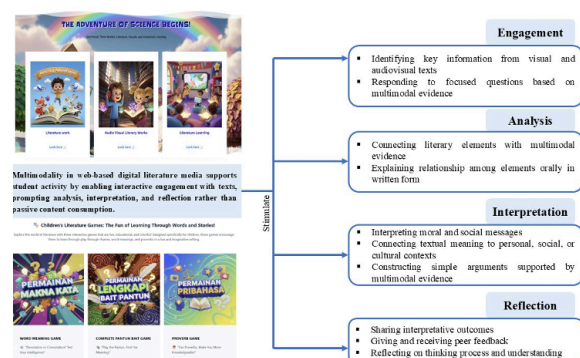
Contemporary literacy practices have shifted from print-dominated texts toward multimodal digital

literacy that integrates verbal, visual, auditory, spatial, and interactive modes (Beltrán-Palanques, 2024; Brosseuk & Downes, 2024; Lim & Kessler, 2024). This transformation necessitates reconceptualizing literacy engagement as a multidimensional construct encompassing cognitive, affective, and behavioral involvement, rather than as passive text consumption. Within the Digital Multimodal Composing (DMC) framework, learners are positioned as active meaning-makers who design and orchestrate multiple semiotic resources to construct meaning, rendering literacy a participatory and reflective process of meaning-making rather than mere skill acquisition (Jiang & Hafner, 2025).

Despite growing interest in multimodal literacy (i.e., the capacity to construct meaning by integrating and orchestrating diverse semiotic resources including images, text, sound, and gesture; see Jewitt, C. and Kress (2003), Mills & Unsworth (2017), and Walsh (2010), recent reviews highlight a persistent gap between digital platform design, instructional practices, and empirically measured literacy engagement outcomes. Much of the existing research remains descriptive and product-oriented, focusing on multimodal artifacts rather than examining how web-based platform features, task structures, and user interactions function pedagogically to foster engagement (J. Lim & Kessler, 2024; Rahmanu & Molnár, 2024). Consequently, there is a need for research that systematically integrates quantitative and qualitative evidence to examine both the effects and mechanisms of engagement within web-based digital literature platforms. Figure 1 presents a pedagogical framework of how multimodal features in a web-based digital literature platform stimulate student engagement. These features activate student-centered activities such as multimodal evidence identification, analytical reasoning, interpretative meaning-making, and reflective interaction.

Figure 1.

Multimodal web-based digital literature media support student engagement by promoting active student participation across engagement, analysis, interpretation, and reflection.



Effectiveness of Digital Platforms and Digital Multimodal Composing

Empirical studies over the past three years suggest that digital platforms and DMC practices can enhance students' literacy engagement and learning outcomes when supported by intentional pedagogical design. Comparative and quasi-experimental research indicates that students involved in multimodal composing demonstrate improved organization, coherence, and rhetorical quality in literacy tasks alongside higher levels of engagement, particularly when reflection on design choices and meaning-making is embedded in instruction (Abdelhalim, 2024; Ho, 2024; Yu et al., 2024).

However, these benefits are not automatic. Without explicit scaffolding, analytical prompting, and structured reflection, digital platforms tend to promote surface-level engagement centered on visual appeal rather than deeper literary analysis (Nurjain, Nurjain, Fajriah, Nurjain, & Firdaus, 2026). This limitation is especially pronounced in elementary contexts, where students require guided support to move from multimodal exposure toward analytical, interpretative, and reflective literacy practices (Susilowaty, Setyarini, & Gustine, 2025). Accordingly, web-based digital literature platforms should be conceptualized not merely as technological tools but as pedagogically mediated instructional environments that deliberately integrate multimodal tasks and progressive scaffolding to support meaningful engagement (Deshmukh & Thakur, 2025).

Pedagogical Design as a Mediator of Multimodal Literacy Engagement

Multimodal literacy theory emphasizes that meaning is constructed through the orchestration of multiple semiotic modes, underscoring the importance of pedagogical designs that make these processes explicit to learners (Brosseau & Downes, 2024; J. Lim & Kessler, 2024). Research consistently demonstrates that literacy engagement increases when students are guided to analyze relationships among visual, textual, and auditory elements and to justify interpretations using multimodal evidence (Lim et al., 2022; Lim & Unsworth, 2023; Midgette et al., 2025). In web-based learning environments, features such as digital annotation, guided prompts, and collaborative discussion spaces are practical only when aligned with clear instructional goals and supported by pedagogical mediation (Ho, 2024).

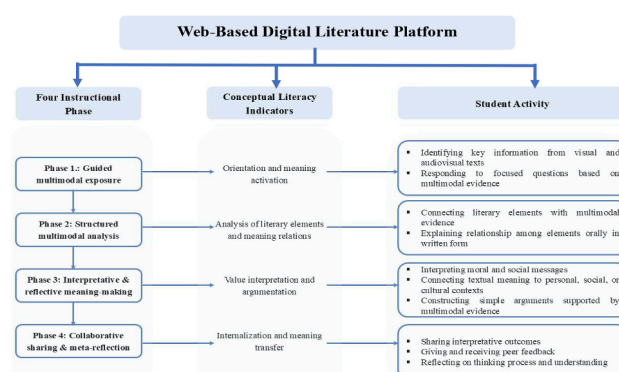
Despite these insights, intervention-based research at the elementary level remains limited, with most empirical studies concentrating on secondary or tertiary education and relying heavily on descriptive or perception-based approaches (Jiang & Hafner, 2025). This gap is particularly consequential given that

elementary education represents a critical phase for the development of literacy engagement, narrative reasoning, and affective responses to multimodal texts (Gacumo et al., 2026; Papen & Peach, 2021). Addressing this limitation, the present study adopts an explanatory sequential mixed-methods design to examine the effects of a pedagogically designed web-based digital literature platform on elementary students' literacy engagement and multimodal literary literacy, while explicating the instructional mechanisms underlying these effects. By integrating outcome-based and process-oriented evidence, this study contributes an applicable instructional framework that clarifies how multimodal digital platforms can be systematically leveraged to foster analytical, interpretative, and reflective literary learning in elementary education (Aloizou, Ioannou, Boloudakis, & Retalis, 2025).

An instructional framework for a web-based digital literature platform illustrating the alignment among four instructional phases, conceptual literacy indicators, and student-centered activities. The framework demonstrates how guided multimodal exposure, structured multimodal analysis, interpretative and reflective meaning-making, and collaborative sharing progressively activate students' cognitive, affective, and behavioral engagement through the identification of multimodal evidence, analytical reasoning, interpretation of moral and social meanings, and reflective interaction (Figure 2).

Figure 2.

Presents the instructional framework developed from the integrated findings of this study. The framework illustrates how multimodal features activate student engagement through four phased student-centered activities: Engagement, Analysis, Interpretation, and Reflection



Method

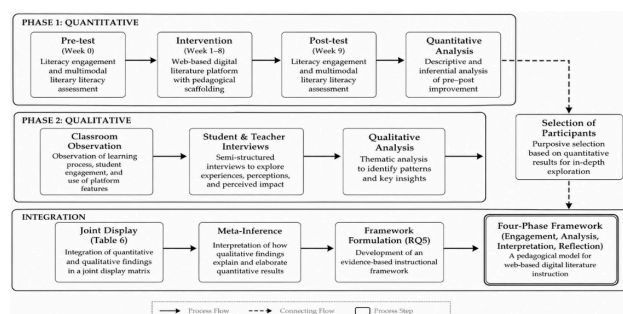
Research Design

This study employed a sequential explanatory mixed-methods design, consisting of two phases: (1) quantitative data collection and analysis to examine the effects of the intervention, followed by (2) qualitative data collection and analysis to explain and

contextualize the quantitative results (Creswell & Clark, 2018). This design was selected for three main reasons. First, the study aimed to measure the effectiveness of the web-based digital literature platform on students' literacy engagement and multimodal literacy literacy, which required quantitative pre- and post-test comparisons to establish statistical evidence of improvement. Second, beyond measuring effectiveness, the study sought to understand how and why the platform produced its effects, which required qualitative exploration of participants' experiences, perceptions, and learning processes. Third, the explanatory sequential design allowed the qualitative findings to elaborate on and explain the quantitative results, providing a more comprehensive and nuanced understanding than either approach could achieve alone (Creswell & Clark, 2018). This design is widely recommended for educational intervention research as it enables researchers to evaluate effectiveness while simultaneously exploring participants' experiences and learning processes (Habibullah et al., 2025).

Figure 3.

Illustrates the sequential phases of the research design, showing the flow from quantitative data collection and analysis to qualitative data collection and analysis, followed by interpretation and integration of both data strands.



Population and Sample

The population comprised fourth-grade students in Bangka Belitung, Indonesia, who were heterogeneous with respect to gender, digital literacy experience, and access to technology. Such heterogeneity was considered a methodological strength, as it enhances ecological validity and reflects real classroom diversity, which has been shown to influence learning outcomes and instructional effectiveness (Decristan, Fauth, Kunter, Büttner, & Klieme, 2017; Gabaldón-Estevan, 2020). Furthermore, in the context of digital literacy, which has been shown to influence elementary students' learning outcomes, a heterogeneous population allows for assessment of how variability in literacy and access affects educational outcomes (Wulandari & Aslam, 2022).

Participants (N = 30) were selected from a fourth-grade classroom at a public elementary school

in Bangka Belitung, Indonesia. The school was selected purposively based on accessibility and the principal's willingness to participate in the eight-week intervention study. The participating class was chosen to reflect typical classroom diversity in terms of gender, digital literacy experience, and access to technology. While this sampling strategy does not permit statistical generalization to the broader population, it was sufficient to examine within-group changes and to generate explanatory qualitative insights that illuminate observed quantitative patterns. The use of a single intact class is common in school-based intervention research where instruction is delivered at the classroom level, as it minimizes instructional contamination and enhances pedagogical authenticity (Cochran, 1977; Kish, 1965). This sampling strategy minimizes instructional contamination and aligns with classroom-based designs that prioritize instructional feasibility and pedagogical authenticity. While not intended for statistical generalization, the sample size was sufficient to examine within-group changes and to generate explanatory qualitative insights that illuminate observed quantitative patterns.

Data collection tools

Classroom observations were conducted to capture key indicators of literacy engagement and instructional practices. The observation protocol focused on five dimensions: (a) student engagement (cognitive, affective, and behavioral indicators), (b) student-teacher interaction, (c) use of digital tools and platform features, (d) collaborative behavior among students, and (e) evidence of multimodal literacy strategies (e.g., use of visual and audio evidence, analytical reasoning).

Observation

Observations were conducted by two trained research assistants who were not involved in the instructional intervention to minimize potential bias. Prior to the observation period, both assistants underwent a four-hour training session that covered: (1) familiarization with the observation protocol and operationalized indicators, (2) practice coding using video recordings of pilot lessons, and (3) discussion of discrepancies to achieve inter-rater reliability. Inter-rater reliability was established with a Cohen's kappa coefficient of 0.85, indicating substantial agreement (Landis & Koch, 1977).

Each classroom session was observed twice weekly over the eight-week intervention period, resulting in a total of 16 observation sessions per class. Each observation lasted for the full duration of the 70-minute lesson. Observations were conducted in real-time using a structured observation sheet with binary scoring (implemented = 1; not implemented = 0) for each indicator, supported by brief descriptive notes to capture contextual details and notable incidents (Chi

& Wylie, 2014; Fredricks et al., 2004; Guimarães & Lima, 2021; Hsiao et al., 2022). To ensure consistency, both observers conducted observations independently during the first two sessions, and any discrepancies were discussed and resolved before continuing with subsequent sessions. The observers were instructed to remain as unobtrusive as possible and to avoid any interaction with students or the teacher during the lessons.

Semi-structured

Semi-structured interviews were employed to explore students’ learning experiences, affective responses, and perceived challenges, as well as teachers’ perspectives on implementation and pedagogical implications. To capture both individual and shared perspectives, two types of interviews were conducted: (1) individual semi-structured interviews with teachers (n = 2) and (2) focus group discussions (FGDs) with students.

The focus group discussions were chosen as the primary method for student data collection for several reasons. First, the interactive nature of focus groups encourages students to build upon each other’s ideas, generating richer and more nuanced insights than individual interviews alone (Krueger & Casey, 2015). Second, since students shared the same classroom experience during the eight-week intervention, the group setting facilitated collective reflection and recall of shared learning events. Third, focus groups are considered developmentally appropriate for elementary students, as they reduce the potential anxiety associated with one-on-one interviews and allow children to feel more comfortable expressing their views in a peer-supported environment (Morgan, 1996)

A total of four focus group discussions were conducted, with each group consisting of five to six students (N = 20). Students were purposively selected to reflect variation in gender, digital literacy experience, and access to technology, ensuring

depth and diversity of qualitative insights (Ahmad & Wilkins, 2025; Jackson et al., 2024). Each FGD session lasted approximately 30-40 minutes and was guided by a semi-structured interview protocol covering students’ engagement experiences, use of platform features, challenges encountered, and perceived learning gains. Additionally, individual interviews were conducted with the two classroom teachers to explore their perspectives on implementation feasibility, pedagogical strategies, and observed changes in student learning. All interviews and FGDs were audio-recorded, transcribed verbatim, and analyzed thematically.

Test

The achievement test was developed to measure students’ multimodal literary literacy, covering three main areas: identifying key information, analyzing intrinsic literary elements, and interpreting moral and social meanings. The test comprised 12 multiple-choice items, which were then validated through expert review and pilot testing as described below. Test items underwent expert review by specialists in literacy and elementary education to establish content validity, followed by minor revisions to improve clarity and communicative appropriateness. Based on expert judgment, 87% of the items (10 out of 12) were deemed ready for use. The remaining two items were identified as having construct validity issues. For instance, one item asked students to determine “what occurs at the conclusion of the tale” with response options that were vague, interpretive, and not mutually exclusive. Such item-writing flaws reduce the ability of items to accurately measure the intended construct (Rezigalla et al., 2024; Sajjad, Iltaf, & Khan, 2020). These items were therefore excluded from the final instrument. Subsequently, an empirical pilot test was conducted with 34 elementary students. Item validity was examined using Pearson’s correlation ($r_{\text{calculated}} > r_{\text{table}}$, $\alpha = .05$, $df = 32$), while internal consistency reliability was assessed using Cronbach’s alpha (Anselmi et al., 2019) in SPSS. Validation results are reported in Table 1.

Table 1.
Validation test results using Pearson correlation

Main Topic	Students' Engagement in Literary Literacy Activities	Item	r _{table}	r _{count}	Sig	Decision
Analyzing information from various types of texts, as well as the values embedded in literary texts derived from visual and/or audiovisual materials	Identifying key information from a literary text	Item 1	.349	.482**	.000	Valid
		Item 2	.349	.471**	.000	Valid
		Item 3	.349	.234*	.023	Invalid
	Determining the story content, theme, plot, setting, and characters of literary texts	Item 4	.349	.370**	.000	Valid
		Item 5	.349	.451**	.000	Valid
		Item 6	.349	.554**	.000	Valid
		Item 7	.349	.369**	.000	Valid
	Articulating the moral, social, or cultural values embedded in literary texts	Item 8	.349	.524**	.000	Valid
		Item 9	.349	.117	.258	Invalid
		Item 10	.349	.553**	.000	Valid
	Concluding the messages or values conveyed in literary texts	Item 11	.349	.440**	.000	Valid
		Item 12	.349	.691**	.000	Valid

Table 1 shows that 10 items were valid and two were invalid. The test demonstrated high reliability (Cronbach's $\alpha = 0.913$). Based on these results, 10 representative items were selected to measure students' engagement in literary activities, with consideration given to balancing test duration with elementary students' attention span to ensure fairness and accurate responses (Chan, Zhang, & Oon, 2023; Rençber & Genç, 2024). The selected items were representative in principle, ensuring that all subtopics and indicators were covered proportionally while avoiding redundancy within any single indicator (Ryoo, Park, Suh, Choi, & Kwon, 2022).

Test structure

Examples of test items that were not selected in the measurement instrument because they were already represented in other items are:

Test number	: 3
Sub-topic	: Information from various types of texts
Students' Engagement in Literary Literacy Activities Indicator	: Identifying key information from a literary text
Item test	: What occurs at the conclusion of the tale Putri Tanjuk Malam? a. She departs for an idyllic and peaceful place. b. The villagers experience profound sorrow. c. The ending remains highly ambiguous and enigmatic. d. All of the provided options may be considered plausible.

The item was determined to be constructively invalid due to several substantive deficiencies. First, the question is ambiguous and insufficiently specific, as it does not direct respondents to any verifiable textual element such as the final event, the protagonist's last action, or a concrete narrative consequence. This lack of specificity prevents the item from accurately eliciting the intended cognitive process. In addition, the answer choices fail to reference explicit textual details; instead, they rely on vague or interpretive descriptors that do not require students to retrieve evidence from the source text. The item also violates the principle of mutually exclusive response options, as option (d) implies that all alternatives may be correct, thereby undermining the item's discriminative capacity and increasing the probability of random guessing. Most critically, the item does not measure the intended construct—identifying key information in a literary text—and instead prompts broad interpretive judgments. Collectively, these flaws indicate that the item lacks construct validity and is therefore unsuitable

for assessing the targeted aspect of literary literacy (Rezigalla et al., 2024). Item-writing flaws lower test quality and reduce the ability of items to differentiate participants based on their mastery of the material (Sajjad et al., 2020).

Data analysis

The normality of the data was examined using the Shapiro–Wilk test. The results showed that both pre-test ($p = .033$) and post-test ($p = .000$) scores violated the normality assumption ($\alpha = .05$). Therefore, the non-parametric Wilcoxon signed-rank test was employed to compare pre- and post-test differences (Goss-Sampson, 2020; Mohammadi Zenouzagh et al., 2025). Effect sizes (Cohen's d) were calculated to estimate the practical magnitude of intervention effects and interpreted using established benchmarks (Lakens, 2022).

Before hypothesis testing, the data were examined for normality using the Shapiro–Wilk test. Given the relatively small sample size ($N = 30$), which is known to increase variability and sensitivity in normality testing (Gosselin, 2024; Souza, Toebe, Mello, & Bittencourt, 2022; Uhm & Yi, 2023), non-normal distributions were anticipated. Therefore, a nonparametric Wilcoxon signed-rank test was employed to examine pre- and post-test differences, as it does not require the assumption of normality (Goss-Sampson, 2020; Mohammadi Zenouzagh et al., 2025). Effect sizes (Cohen's d) were calculated to estimate the practical magnitude of intervention effects and interpreted using established benchmarks (Lakens, 2022).

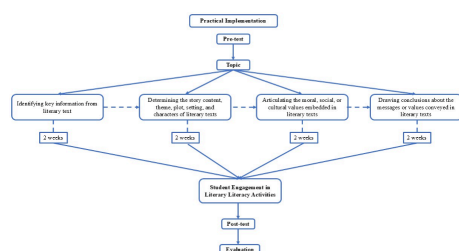
The qualitative phase aimed to explain quantitative outcomes. Structured classroom observations documented student engagement, teacher–student interaction, use of digital tools, and enactment of multimodal literacy strategies. Semi-structured interviews with purposively selected students and teachers explored learning experiences, perceived challenges, and meaning-making processes. Qualitative data were analyzed using thematic analysis and triangulated with quantitative results to strengthen explanatory validity.

Finally, interpretive integration combined statistical findings and qualitative themes to address the research questions and inform the development of an applicable instructional framework, consistent with explanatory sequential mixed-methods principles in educational research (Rezigalla et al., 2024).

Experimental Process

This study employed a mixed-methods sequential explanatory design, as illustrated in Figure 3.

Figure 3.
Experimental Process



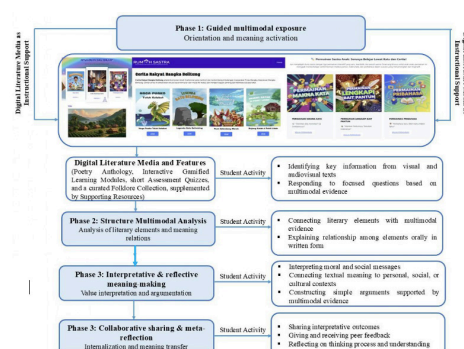
The experimental procedure began with a pre-test to assess students' baseline literacy before the intervention (Amumpuni, Hartono, Rukmini, & Haryanti, 2025). The instructional intervention was then implemented over eight weeks, structured into four sequential instructional phases, each lasting two weeks. These phases were designed to progressively develop core literary competencies, beginning with identifying key information in scholarly texts and culminating in concluding thematic, moral, and social meanings embedded in the texts (Kong, 2014; Puspitasari, Widiarti, & Subali, 2025)

The intervention emphasized both analytical processes—including identifying story content, theme, plot, setting, and characters—and interpretative processes, such as articulating moral, social, and cultural values, with the explicit aim of strengthening students' cognitive engagement in literary activities. Following the intervention, a post-test was administered to assess students' post-intervention literacy performance. Pre- and post-test data were subsequently analyzed to determine the significance of the intervention's learning gains.

Teaching Environment to Engage Students

The learning environment implemented in this study reflects a contemporary instructional approach that integrates web-based digital literature media to support multimodal, contextual, and participatory literary literacy, as illustrated in Figure 4.

Figure 4.
Environment to engage students



Digital media function not merely as content repositories but as instructional supports that provide access to literary texts, images, audio, animations, and interactive quizzes, thereby creating a multimodal, resource-rich, flexible, and accessible learning environment. The platform facilitates interaction between students and literary content, as well as teacher-student and peer interaction through gamified modules, quizzes, and collaborative tasks, which collectively enhance student engagement, motivation, and active participation (As Syafaatussalamah & Deandra Eka Salsabilla, 2025; Hardiyanti, Fitriani, & Ashari, 2023; Mohammadi Zenouzagh et al., 2025; Yuliana, Nirmala, & Ardiasih, 2023).

By incorporating locally grounded literary texts, the learning environment situates literary literacy within students' cultural contexts while leveraging digital affordances. This integrated environment—combining digital space, scholarly content, interaction, technology, and cultural relevance—demonstrates strong potential for enhancing student engagement and learning outcomes in the digital literacy era.

Result

This section presents the findings organized according to the five research questions. The results are reported sequentially, beginning with quantitative outcomes examining the effectiveness of the intervention (RQ1 and RQ2), followed by qualitative findings that explain the processes underlying these outcomes (RQ3 and RQ4), and concluding with the integration of both data strands to formulate an evidence-based instructional framework (RQ5). Quantitative data were analyzed using descriptive statistics and the Wilcoxon signed-rank test, while qualitative data from classroom observations, student interviews, and teacher interviews were analyzed.

Effect of the Web-Based Digital Literature Platform on Students' Literacy Engagement

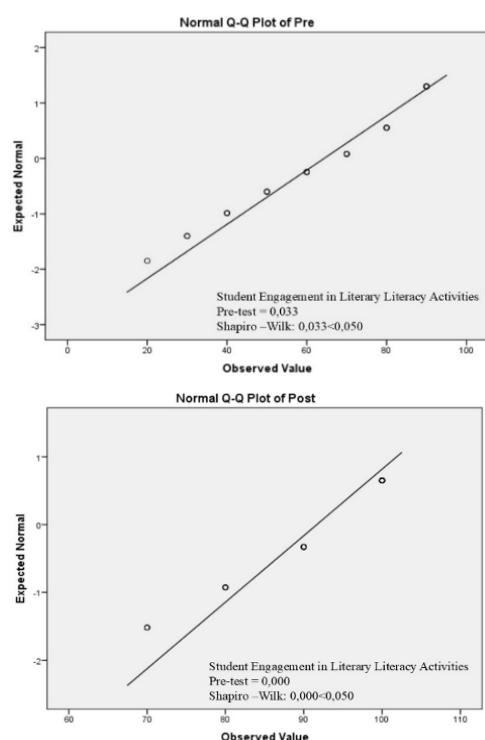
Before hypothesis testing, data were examined for normality using the Shapiro-Wilk test. The results indicated that student engagement scores were not normally distributed, with p-values of .033 for the pre-test and .000 for the post-test (Figure 5). As both values were below the $\alpha = .05$ threshold, the assumption of normality was not met.

The Wilcoxon test revealed a statistically significant increase in students' engagement in literary activities following the eight-week intervention using the web-based digital literature platform ($p < .001$). As shown in Table 2, the mean engagement score increased markedly from 64.33 (SD = 20.46) in the pre-test to 91.67 (SD = 10.20) in the post-test. The calculated effect size was considerable (Cohen's $d_z = 1.69$), substantially

exceeding the conventional benchmark for a significant effect ($d = 0.80$). This finding indicates that the intervention produced a strong and educationally meaningful impact on students' literacy engagement.

Figure 5.

Q-Q Plot of Pre-Post Test



Collectively, these results demonstrate that the pedagogically designed web-based digital literature platform significantly enhanced students' literacy engagement across cognitive, affective, and behavioral dimensions compared to their pre-intervention levels.

Perceptions and Responses in Multimodal Literacy Skills

To address RQ2, Table 3 synthesizes quantitative pre-post assessment outcomes and qualitative evidence from observations and interviews, providing a triangulated account of students' development in identifying key information, analyzing intrinsic elements, and interpreting moral and social meanings in multimodal literary texts.

The analysis demonstrates convergence across quantitative and qualitative data sources, indicating that assessment results, classroom observations, and participant perspectives consistently supported improvements in students' multimodal literacy skills. This triangulated evidence strengthens the validity of the findings related to RQ2 and provides a robust basis for subsequent integration with RQ1 and the development of the instructional framework (RQ5).

Students' and Teachers' Experiences and Perceptions of Literacy Engagement and Literary Analysis

To address RQ3, qualitative data from classroom observations and interviews with students and teachers were analyzed to explain the quantitative gains in literacy engagement (RQ1) and multimodal literacy skills (RQ2). Using an explanatory sequential mixed-methods approach, a thematic matrix was employed to triangulate data sources and elucidate how the web-based digital literature platform shaped students' engagement, analytical reasoning, and classroom literacy practices (Table 4).

As summarized in Table 4, students' and teachers' experiences provide a coherent explanatory account of the quantitative improvements observed in literacy engagement and multimodal literary literacy skills. Increased engagement functioned as a foundational mechanism that enabled focused attention, analytical processing, and reflective interpretation of multimodal literary texts. For instance, one student noted that "the pictures and animations make me want to follow the story until the end" (S2), while another explained that "the pictures help me understand what the story is about quickly" (S5). A teacher similarly confirmed that "students are significantly more attentive when digital media is used" (T1). The convergence of observational data and participant perspectives demonstrates that the web-based digital literature platform did not merely enhance motivation but systematically reshaped students' cognitive, affective, and interpretative literacy practices. Another student reflected on the interpretive dimension, stating that "the story reminds me of traditions in our village" (S7), indicating how the platform facilitated connections between literary texts and students' cultural contexts. A teacher also observed that "visuals helped students identify essential information more confidently" (T2), underscoring the platform's role in supporting analytical competence. These quotations illustrate how the platform transformed students from passive viewers to active meaning-makers, supporting the quantitative gains reported earlier.

Use of Multimodal Features to Build Literary Understanding

Building on the significant gains in literacy engagement (RQ1), multimodal literary literacy skills (RQ2), and students' and teachers' reported experiences (RQ3), this section examines how students actively utilized the multimodal features and pedagogical tasks embedded in the web-based digital literature platform. Qualitative evidence from classroom observations and interviews reveals that students' analytical, interpretative, and reflective understanding developed through purposeful interaction with specific platform affordances rather than through exposure to digital media alone.

Table 2.
The Result of the t-test through Wilcoxon

	N	Mean	Std. Deviation	Minimum	Maximum	t-test	Cohen's Effect
Pre	30	64.333	20.457	20.000	90.000	0.000	1.690
Post	30	91.667	10.1992	70.000	100.000		

Table 3.
Multimodal Literary Literacy Skills across Data Sources

RQ2 Indicator	Quantitative Evidence (Pre-Post Assessment)	Observation Data	Student Interview Excerpts	Teacher Interview Insights
Identifying key information	Increased accuracy in identifying central ideas and relevant events in post-test responses	Students focused on essential scenes rather than surface-level visuals	"Now I know which parts of the story are important."	"Students grasp the core ideas more quickly."
Analyzing intrinsic elements (theme, plot, setting)	Post-test responses demonstrated a more transparent structure and more logical sequencing	Students explained plot development and character actions coherently	"The plot is clearer when there are images and videos."	"They have started linking themes to events."
Interpreting moral and social meanings	Written responses showed greater depth and more vigorous justification	Students referred to visual and audiovisual cues during discussions	"The values are visible through the characters' actions."	"Students' answers are no longer limited to naming values."
Use of multimodal evidence	More frequent integration of visual and audio references	Students pointed to images, sounds, and scenes as supporting evidence	"The music helps me understand the mood."	"Students are beginning to use evidence from videos."
Level of literary reasoning	A shift from descriptive to analytical responses	Increased participation in interpretive discussions	"I can explain my reasons now."	"Classroom discussions have become more dynamic."

Table 4.
Students' and Teachers' Experiences and Perceptions of Literacy Engagement and Literary Analysis

Core Theme	Analytical Focus (Link to RQ1-RQ2)	Evidence from Classroom Observation	Evidence from Student Interviews	Evidence from Teacher Interviews
Increased engagement in multimodal literary activities	Engagement as a prerequisite for learning gains (RQ1)	Sustained attention during audiovisual sessions; reduced off-task behavior; increased participation	"The pictures and animations make me want to follow the story until the end."	"Students are significantly more attentive when digital media is used."
Improved identification of key information	Transition from engagement to cognitive processing (RQ2a)	Students used visual cues to identify main events more efficiently	"The pictures help me understand what the story is about quickly."	"Visuals helped students identify essential information more confidently."
Enhanced comprehension of literary elements (theme, plot, setting)	Development of analytical literary reasoning (RQ2b)	Coherent sequencing of plot; justified identification of themes using visual/audio cues	"I can understand the story because I see what happens next."	"Media provides contextual cues for deeper literary understanding."
Recognition of moral, social, and cultural values	Shift from analysis to reflective interpretation (RQ2c)	Students connected story events to local cultural contexts (Bangka Belitung)	"The story reminds me of traditions in our village."	"Local narratives made values feel personal and relevant."
Increased ability to draw logical conclusions	Emergence of evidence-based reasoning (RQ2c)	Students justified their interpretations using images and narrative audio	"I can show the part of the video that proves my answer."	"Visual evidence supports logical argumentation."
Positive affective response and learner confidence	Affective dimension sustaining engagement (RQ1)	Increased enthusiasm; voluntary participation; collaborative discussion	"I'm not afraid to give my opinion."	"Previously passive students became more confident."
Pedagogical affordances for teachers	Instructional mediation and scaffolding	Structured prompts guided discussions	—	"The platform reduces explanation time and improves clarity."
Technical and implementation challenges	Critical reflection on feasibility and scalability	Occasional connectivity issues; need for navigation support	"Sometimes the video loads slowly."	"Limited devices slow group progress."

As summarized in Table 5, students' analytical, interpretative, and reflective understanding emerged through purposeful interaction with specific multimodal features and pedagogical tasks. These affordances served as instructional scaffolds, transforming increased engagement (RQ1) into higher-order multimodal literary literacy (RQ2), as further explained by students' and teachers' experiences (RQ3).

Integration of Quantitative and Qualitative Findings

Table 6 presents the integrated findings addressing RQ5 through triangulation and thematic analysis.

The table functions as a joint display that synthesizes quantitative evidence (RQ1–RQ2) and qualitative insights (RQ3–RQ4) by aligning literacy stages, instructional phases, instructional inputs and designs, emergent thematic categories, and empirical evidence. This integrative mapping explicates the pedagogical mechanisms underlying students' engagement, analysis, interpretation, and reflection, thereby informing the development of an evidence-based instructional framework for web-based multimodal digital literature learning in elementary education.

Table 5.

Analysis of Use of Multimodal Features to Build Literary Understanding

Multimodal Feature / Task	Learning Function	Evidence from Classroom Observation	Evidence from Student Interviews	Evidence from Teacher Interviews
Audiovisual narratives (animation, audio)	Supporting narrative comprehension and plot analysis	Students followed the story flow more accurately and discussed events coherently	"The pictures and sound help me understand the story better."	"Students grasp the storyline more easily with audiovisual support."
Pause and replay functions.	Self-paced analytical processing	Students revisited key scenes to identify essential events	"I can watch important parts again."	"Replay helps students analyze details."
Guided analytical questions	Scaffolding analytical reasoning	Students structured their responses around prompts	"The questions help me think about the story."	"Guiding questions make analysis more focused."
Interactive multimodal tasks (group work)	Collaborative meaning-making	Increased peer discussion and shared interpretation	"It's easier to discuss when we see the pictures together."	"Group tasks increase participation."
Multimodal evidence prompts	Encouraging evidence-based interpretation	Students cited images and sounds to justify responses	"I can show the picture that proves my answer."	"Students now use visual evidence."
Reflective value-based tasks	Fostering moral and social reflection	Students linked story values to personal experiences	"The story teaches us to help others."	"Reflection tasks deepen ethical discussion."

Table 6.

Analysis of Integration of Quantitative and Qualitative Findings

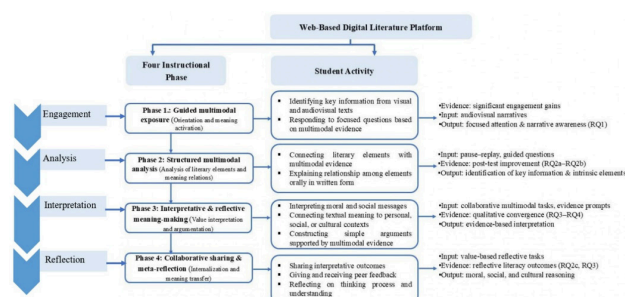
Literacy Stage	Instructional Phase	Instructional Input	Instructional Design	Cognitive & Literacy Output	Thematic Analysis	Empirical Evidence (RQ)	Pedagogical Function (Meta-Inference)
Engagement	Phase 1: Guided Multimodal Exposure	Audiovisual narratives (cerita digital multimodal)	Pertanyaan terpadu, eksplorasi visual-audio	Focused attention, narrative awareness	Meaning activation	RQ1 (significant engagement gains)	Affective activation and meaning schemata as the foundation of literacy
Analysis	Phase 2: Structured / Focused Meaning Construction	Pause-replay, guided questions	Worksheet multimodal, diskusi terstruktur	Identifikasi informasi kunci & unsur intrinsik	Analytical literacy	RQ2a–RQ2b (post-test improvement)	Multimodal evidence-based literary analysis scaffolding
Interpretation	Phase 3: Interpretative Meaning-Making	Collaborative multimodal tasks, evidence prompts	Diskusi reflektif, koneksi nilai	Interpretasi berbasis bukti multimodal	Value interpretation	RQ3–RQ4 (qualitative convergence)	Meaning construction, justification, and argumentation
Reflection	Phase 4: Collaborative Reflection / Value-Oriented Response	Value-based reflective tasks	Peer feedback, meta-reflection	Penalaran moral, sosial, dan budaya	Metacognitive internalization	RQ2c, RQ3 (reflective literacy outcomes)	Internalization of values and transfer of understanding

Based on the triangulation of quantitative results (RQ1–RQ2) and qualitative findings (RQ3–RQ4) through thematic analysis, this study formulates a web-based multimodal digital literature instructional framework operating across four progressive literacy phases: engagement, analysis, interpretation, and reflection.

These phases do not function in isolation; instead, they constitute a sequenced pedagogical mechanism that explains how learning gains are generated. The guided multimodal exposure phase activates initial meaning-making and student engagement (RQ1), which serve as prerequisites for evidence-based literary analysis in the structured analysis phase (RQ2a–RQ2b). Subsequently, the interpretative meaning-making phase enables students to collaboratively construct literary meanings and values, as indicated by the convergence of qualitative evidence (RQ3–RQ4). This process culminates in the collaborative reflection phase, where metacognitive reflection and value-based reasoning signal the internalization and transfer of meaning (RQ2c, RQ3).

Accordingly, RQ5 is addressed through meta-inference, demonstrating that effective multimodal digital literature instruction in elementary education requires a staged, evidence-informed, and reflective instructional design, rather than the mere adoption of digital technologies or texts. The conceptual framework for web-based multimodal digital literature instruction is presented in Figure 6.

Figure 6.
The conceptual framework for web-based multimodal digital literature instruction is presented



Discussion

The findings addressing RQ1 suggest that the implementation of the pedagogically designed web-based digital literature platform created learning conditions that encouraged students to become more actively involved in literacy activities. Classroom observations indicated that students demonstrated greater persistence in exploring stories, revisiting narrative segments, and participating in discussions compared to their engagement during conventional literacy instruction. This pattern suggests that literacy engagement in this study was not merely reflected in students' enjoyment of digital media but in their

increased willingness to invest attention and cognitive effort in understanding literary texts.

The findings addressing RQ1 demonstrate that the eight-week intervention using a pedagogically designed web-based digital literature platform produced a substantial and educationally meaningful improvement in students' literacy engagement. The magnitude of this effect suggests that the intervention was practically significant in its impact on student learning. The findings addressing RQ1 demonstrate that the eight-week intervention produced a substantial and educationally meaningful improvement in students' literacy engagement. This large effect size aligns with contemporary research positioning multimodal text environments as catalysts for student motivation, focused attention, and active participation, which are widely recognized as prerequisites for deeper cognitive processing. However, the present study indicates that the positive impact of multimodal environments cannot be attributed solely to technological features. Field observations revealed that students' engagement became more meaningful when multimodal resources were accompanied by guided learning activities that directed attention toward story content and interpretation. Therefore, engagement should be understood not merely as an affective outcome but as a pedagogical mechanism that enables students to access textual meaning and participate more deeply in literacy learning processes.

The triangulated findings addressing RQ2 demonstrate that students developed stronger multimodal literary literacy, particularly in identifying key information, analyzing intrinsic literary elements, and interpreting moral and social meanings embedded in literary texts. Qualitative evidence revealed that students gradually progressed from retelling story events toward providing explanations and interpretations supported by textual, visual, and auditory evidence. Such development suggests that multimodal literacy involves more than comprehension of narrative content; it requires learners to integrate multiple semiotic resources in constructing meaning.

These findings are consistent with studies showing that structured multimodal learning activities, including guided questioning, replay opportunities, and evidence-based tasks, can enhance learners' internal representations and facilitate analytical reasoning (Liang & Hwang, 2023; Shen et al., 2023). Nevertheless, the present study extends previous findings by demonstrating that such cognitive development can occur effectively in elementary education contexts when multimodal affordances are systematically integrated with instructional scaffolding. This suggests that multimodal literacy growth is not an automatic consequence of exposure to digital texts but rather the result of carefully designed pedagogical

processes that support evidence gathering, analysis, and interpretation.

The qualitative findings addressing RQ3 further illuminate the mechanisms through which literacy development occurred. Interviews with students and teachers revealed that audiovisual features facilitated access to narrative contexts and reduced barriers to understanding literary texts. This finding is particularly relevant in elementary classrooms, where students often experience difficulty visualizing narrative settings, character relationships, and plot developments from written text alone. The availability of multimodal representations appeared to provide concrete references that supported comprehension and interpretation.

However, the findings also revealed that multimodal resources alone were insufficient to promote deeper literary understanding. Observations consistently showed that students rarely moved beyond surface-level observations without teacher intervention. Through guided questioning, interpretive modeling, and collaborative discussion, teachers encouraged students to examine narrative evidence, justify interpretations, and consider alternative perspectives. This finding supports theoretical arguments that multimodal pedagogy becomes educationally meaningful only when interactions among learners, texts, and multimodal resources are mediated through purposeful instructional design (Linder & Falk-Ross, 2024). It also aligns with sociocultural perspectives that view literacy learning as a socially mediated process in which meaning is co-constructed through interaction and guided participation.

The findings addressing RQ4 provide further insight into the pedagogical functions of specific platform features and learning activities. Observational and interview data suggest that features such as pause-replay tools, guided prompts, collaborative tasks, evidence prompts, and reflective activities played distinct but interconnected roles in supporting literacy development. Replay functions enabled students to revisit important scenes and verify interpretations using visual and auditory evidence. Guided prompts helped structure students' attention toward significant narrative elements, while collaborative activities created opportunities for negotiating and refining interpretations through peer interaction.

These findings corroborate research emphasizing that effective multimodal learning is inherently interactive and mediated rather than passive or technology-driven (Liang & Hwang, 2023; Shen et al., 2023). More importantly, the present study demonstrates how specific instructional supports transform multimodal exposure into meaningful learning experiences. The findings suggest that the educational value of multimodal platforms lies not

simply in presenting information through multiple modes but in creating opportunities for learners to engage critically with evidence, justify interpretations, and reflect on meaning. This reinforces constructivist perspectives that knowledge develops through active engagement, social interaction, and reflective inquiry.

The integration of findings across RQ1–RQ4 through triangulation and thematic analysis resulted in the identification of a four-phase instructional framework consisting of Engagement, Analysis, Interpretation, and Reflection. Rather than functioning as a descriptive sequence of learning activities, this framework provides an explanatory model of the pedagogical mechanisms underlying students' literacy development. Engagement establishes the motivational and attentional foundation necessary for participation. Analysis directs learners toward the examination of multimodal evidence. Interpretation enables students to construct and justify meaning collaboratively, while Reflection supports the internalization of values and the transfer of understanding beyond immediate learning tasks.

This framework aligns with systematic reviews of multimodal pedagogy emphasizing the importance of instructional sequencing, scaffolding, and opportunities for meaning-making across modes (Linder & Falk-Ross, 2024). At the same time, the present findings contribute a more process-oriented explanation of how literacy development unfolds in elementary digital literature instruction. The framework suggests that engagement alone is insufficient to generate meaningful learning outcomes unless it is connected to analytical, interpretive, and reflective processes. Consequently, literacy development should be understood as a progressive and interconnected process rather than a collection of isolated skills.

From a theoretical perspective, this study contributes to the growing discourse on multiliteracies and multimodal pedagogy by providing empirical evidence from an elementary education context. The findings support the proposition that literacy in contemporary digital environments extends beyond conventional print-based comprehension and involves the capacity to interpret, evaluate, and construct meaning across multiple modes of communication (Linder & Falk-Ross, 2024; Shen et al., 2023). More importantly, this study advances current understanding by demonstrating that multimodal literacy development emerges through the interaction between technological affordances and pedagogically designed learning experiences. Thus, the findings challenge technology-deterministic assumptions and reinforce the view that instructional design remains central to meaningful literacy learning.

From a practical perspective, the findings offer important implications for elementary literacy

education. In many educational settings, digital technologies are often introduced with the expectation that technological innovation alone will improve learning outcomes. However, the present findings indicate that educational benefits emerge when digital resources are integrated with structured scaffolding, collaborative inquiry, evidence-based reasoning, and reflective learning activities. Therefore, educators should focus not only on adopting digital tools but also on designing pedagogical experiences that guide students toward deeper engagement and interpretation.

Although this study was conducted in an elementary school context in Bangka Belitung using locally grounded digital literary texts, the resulting instructional framework emphasizes pedagogical mechanisms rather than cultural specificity. Guided multimodal scaffolding, analytical prompting, collaborative interpretation, and reflective meaning-making represent transferable principles that may be adapted across diverse educational contexts where digital and multimodal texts are employed to support literacy learning.

Several limitations should be acknowledged. The relatively small sample size and the limited duration of the intervention restrict the extent to which findings can be generalized across broader populations and longer periods of implementation. Furthermore, challenges related to internet connectivity and device availability highlight practical constraints that may influence scalability, particularly in resource-diverse educational environments. Similar concerns have been reported in previous educational technology studies (Liang & Hwang, 2023). Future research should therefore investigate longitudinal impacts, examine implementation across varied educational settings, and explore differentiated scaffolding strategies tailored to students with diverse literacy profiles.

Overall, this study demonstrates that web-based digital literature platforms can support meaningful literacy development when embedded within coherent pedagogical structures. The findings indicate that literacy growth occurs through the interaction of multimodal resources, instructional scaffolding, collaborative meaning-making, and reflective engagement. The resulting four-phase instructional framework offers both a theoretical contribution to multimodal literacy research and a practical guide for educators seeking to design, implement, and evaluate effective digital literature instruction in elementary education.

Conclusion

This study demonstrates that a pedagogically designed web-based digital literature platform can meaningfully enhance elementary students'

engagement with literacy and multimodal literacy when technology is embedded within a structured instructional design. The explanatory sequential mixed-methods findings indicate that improvements in engagement functioned as an enabling condition for deeper analytical, interpretative, and reflective literary processing, particularly when supported by guided multimodal scaffolding, analytical prompts, and evidence-based tasks. The resulting four-phase instructional framework—Engagement, Analysis, Interpretation, and Reflection—extends existing multimodal literacy research by explicating the pedagogical mechanisms through which digital platforms support literary meaning-making, positioning the local classroom as a pedagogical laboratory for innovation rather than a contextual limitation.

Nevertheless, several limitations warrant consideration. The relatively small sample size and the eight-week duration constrain the generalizability and longitudinal interpretation of the findings. In addition, variations in device availability and connectivity highlight implementation challenges that may affect scalability across diverse educational settings. The study primarily focused on a single instructional configuration, leaving open questions about differentiated scaffolding for learners with varied literacy profiles.

Based on the findings, the following evidence-based recommendations are offered for teachers and decision-makers.

For Teachers. Teachers should integrate guided questioning and interpretive modeling to transform students' surface-level viewing into analytical reasoning. They are encouraged to utilize pause-replay features and guided prompts to support evidence-based justification and to follow the four-phase instructional sequence (Engagement, Analysis, Interpretation, Reflection) to progressively scaffold literacy development. Additional scaffolding should be provided for students with lower initial literacy levels, as these students showed the most significant gains with structured multimodal support.

For Decision-Makers. Professional development should prioritize multimodal pedagogical mediation over technical training, as teacher scaffolding was critical to successful implementation. Digital infrastructure and device availability must be improved, as connectivity and access barriers hindered equitable participation. Curriculum frameworks should explicitly incorporate multimodal literary competencies and the four-phase instructional framework. Sustained investment is needed to support multi-site implementations and assess the framework's transferability across diverse contexts.

For Future Research. Longitudinal and multi-site studies are needed to test the framework's durability and transferability across contexts and learner populations. Further research should also explore differentiated scaffolding strategies for students with varied literacy profiles and develop infrastructure-adaptive designs for resource-limited settings.

Ethics Statements:

This research has adhered to the ethical standards of this journal.

Conflict of Interest:

The author declares that there is no conflict of interest in this publication.

Informed Consent:

All participants have agreed to participate in this study, as stated in the informed consent form, with their real names removed.

Data Availability:

All the data collected from the participants' test and observation results.

References

- Abdelhalim, S. M. (2024). From traditional writing to digital multimodal composing: promoting high school EFL students' writing self-regulation and self-efficacy. *Computer Assisted Language Learning*, 1–30. doi: 10.1080/09588221.2024.2322148
- Ahmad, M., & Wilkins, S. (2025). Purposive sampling in qualitative research: a framework for the entire journey. *Quality & Quantity*, 59(2), 1461–1479. doi: 10.1007/s11135-024-02022-5
- Aloizou, V., Ioannou, A., Boloudakis, M., & Retalis, S. (2025). A learning experience design framework for multimodal learning in the early childhood. *Smart Learning Environments*, 12(1), 27. doi: 10.1186/s40561-025-00376-3
- Amumpuni, R. S., Hartono, R., Rukmini, D., & Haryanti, R. P. (2025). The impact of collaborative strategic reading on EFL students with different critical thinking levels. *Social Sciences & Humanities Open*, 12, 101904. doi: 10.1016/j.ssaho.2025.101904
- Anselmi, P., Colledani, D., & Robusto, E. (2019). A Comparison of Classical and Modern Measures of Internal Consistency. *Frontiers in Psychology*, 10. doi: 10.3389/fpsyg.2019.02714
- As Syafaatussalamah, & Deandra Eka Salsabilla. (2025). Efektivitas penggunaan media digital dalam meningkatkan motivasi belajar siswa sekolah dasar. *Al-Tarbiyah : Jurnal Ilmu Pendidikan Islam*, 3(3), 11–24. doi: 10.59059/al-tarbiyah.v3i3.2479
- Baker, M. J., Pallarès, G., Cedar, T., Brandel, N., Biatti, L., Schwarz, B., & Détienne, F. (2023). Understanding the moral of the story: Collaborative interpretation of visual narratives. *Learning, Culture and Social Interaction*, 39, 100700. doi: 10.1016/j.lcsi.2023.100700
- Beltrán-Palanques, V. (2024). Assessing video game narratives: Implications for the assessment of multimodal literacy in ESP. *Assessing Writing*, 60, 100809. doi: 10.1016/j.asw.2024.100809
- Brosseuk, D., & Downes, L. (2024). Listening to teachers talk about multimodality and multimodal texts: considerations for the national English curriculum. *Australian Journal of Language and Literacy*, 47(3), 317–334. doi: 10.1007/s44020-024-00064-8
- Calafato, R., & Simmonds, K. (2022). Linkages between literary response, aesthetic competence, and literary competence in the EFL classroom. *Studies in Educational Evaluation*, 75(August), 101214. doi: 10.1016/j.stueduc.2022.101214
- Chan, W. K., Zhang, L., & Oon, E. (2023). Classroom assessment that tailor instruction and direct learning: A validation study. *International Journal of Assessment Tools in Education*, 10(2), 376–394. doi: 10.21449/ijate.1155679
- Chi, M. T. H., & Wylie, R. (2014). The ICAP Framework: Linking cognitive engagement to active learning outcomes. *Educational Psychologist*, 49(4), 219–243. doi: 10.1080/00461520.2014.965823
- Chourio-Acevedo, L., Köhler, J., Coscarelli, C., Gacitúa, D., Proaño-Ríos, V., & González-Ibáñez, R. (2024). Information literacy development and assessment at school level: a systematic review of the literature. *ArXiv*, 2404.19020. doi: https://doi.org/10.48550/arXiv.2404.19020
- Cochran, W. G. (1977). *Sampling Techniques* (3rd ed.). New York: John Wiley & Sons.
- Creswell, J. W., & Clark, P. (2018). *Designing and conducting mixed methods research*. New York: Sage Publication, Inc.
- Creswell, J. W., & Creswell, J. D. (2018). Mixed Methods Procedures. In *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches*. California: Sage Publication, Inc.

- Dahiru, M. (2020). Challenges of teaching literature-in-english in secondary schools in Borno and Yobe, Nigeria. *International Journal of Linguistics, Literature and Translation*, 3(12), 233–244. doi: 10.32996/ijllt.2020.3.12.27
- Decristan, J., Fauth, B., Kunter, M., Büttner, G., & Klieme, E. (2017). The interplay between class heterogeneity and teaching quality in primary school. *International Journal of Educational Research*, 86, 109–121. doi: 10.1016/j.ijer.2017.09.004
- Deshmukh, N. D., & Thakur, S. A. (2025). Integrating multimodal teaching strategies into STEM Education: Reflections from pre-service and in-service teacher workshop experiences. *science and technology education: Expectations and experiences*, 71–85. Scientia Socialis Press. doi: 10.33225/BalticSTE/2025.71
- Dey, A., & Munshi, S. A. (2025). Fun with images: An analysis of the role of visual literacy in facilitating easy and enjoyable learning with a focus on future prospects. *Libri*, 75(2), 109–127. doi: 10.1515/libri-2024-0143
- Encabo-Fernández, E., Albarracín-Vivo, D., & Jerez-Martínez, I. (2023). Evaluative research on the critical thinking of primary school students. *International Journal of Educational Research Open*, 4, 100249. doi: 10.1016/j.ijedro.2023.100249
- Eryilmaz, M. A., & Çelik, Ö. C. (2025). Visual literacy in the primary school Turkish Curriculum: Insights from Classroom Teachers. *Ana Dili Eğitimi Dergisi*, 13(4), 883–902. doi: 10.16916/aded.1721329
- Farrar, J., Arizpe, E., & Lees, R. (2024). Thinking and learning through images: a review of research related to visual literacy, children's reading and children's literature. *Education 3-13*, 52(7), 993–1005. doi: 10.1080/03004279.2024.2357892
- Firdaus, S., Komarudin, & Gloriani, Y. (2024). The effectiveness of use of multimodal text teaching materials on learning the Indonesian language in high school. *International Journal of Educational Research Excellence (IJERE)*, 3(2), 657–667. doi: 10.55299/ijere.v3i2.970
- Fradana, A. N., Suwarta, N., & Sotlikova, R. (2025). Strengthening multimodal literacy through digital literary text learning: Enhancing students' achievement in the age of disruption. *Journal of Language and Literature Studies*, 5(1), 216–225. doi: 10.36312/jolls.v5i1.2635
- Fredricks, J. A., Blumenfeld, P. C., & Paris, A. H. (2004). School Engagement: Potential of the Concept, State of the Evidence. *Review of Educational Research*, 74(1), 59–109. doi: 10.3102/00346543074001059
- Gabaldón-Estevan, D. (2020). Heterogeneity versus Homogeneity in Schools: A Study of the Educational Value of Classroom Interaction. *Education Sciences*, 10(11), 335. doi: 10.3390/educsci10110335
- Gacumo, R. J., Coats, K., & Kucirkova, N. I. (2026). Reading without words: sensory inclusion and meaning-making in children's multimodal texts. *Children's Literature in Education*. doi: 10.1007/s10583-026-09677-8
- Ganapathy, M., & Seetharam, S. A. (2016). The Effects of Using Multimodal Approaches in Meaning-Making of 21st Century Literacy Texts Among ESL Students in a Private School in Malaysia. *Advances in Language and Literary Studies*, 7(2). doi: 10.7575/aiac.all.v.7n.2p.143
- Gasser, L., Preisig, D., Frei, A., Dammert, Y., Egger, S., & Murphy, P. K. (2025). Effects of a literary intervention on interaction quality in small-group discussions in the upper elementary grades. *Learning and Instruction*, 95(September 2024), 101996. doi: 10.1016/j.learninstruc.2024.101996
- Gil-Flores, J., Rodríguez-Santero, J., & Torres-Gordillo, J.-J. (2017). Factors that explain the use of ICT in secondary-education classrooms: The role of teacher characteristics and school infrastructure. *Computers in Human Behavior*, 68, 441–449. doi: 10.1016/j.chb.2016.11.057
- Gnams, T. (2025). Reciprocal effects between information and communication technology literacy and conventional literacies. *Intelligence*, 111(February), 101936. doi: 10.1016/j.intell.2025.101936
- Goss-Sampson, M. A. (2020). Statistical Analysis in JASP A Guide for Students. In JASP. JASP.
- Gosselin, R.-D. (2024). Testing for normality: a user's (cautionary) guide. *Laboratory Animals*, 58(5), 433–437. doi: 10.1177/00236772241276808
- Guimarães, L. M., & Lima, R. da S. (2021). A systematic literature review of classroom observation protocols and their adequacy for engineering education in active learning environments. *European Journal of Engineering Education*, 46(6), 908–930. doi: 10.1080/03043797.2021.1937946

- Habibullah, J. A., Norvaizi, I., & Dewi, D. E. C. (2025). Implementasi Mixed Methods dalam Penelitian Pendidikan. *Peradaban Journal of Interdisciplinary Educational Research*, 3(1), 17–31. doi: 10.59001/pjier.v3i1.245
- Hardiyanti, E., Fitriani, V., & Ashari, W. (2023). The Influence of Digital Tools on Student Engagement and Academic Outcomes Across Educational Levels. *Educia Journal*, 1(2), 59–68. doi: 10.71435/610408
- Hew, K. F., & Brush, T. (2007). Integrating Technology into K-12 Teaching and Learning: Current Knowledge Gaps and Recommendations for Future Research. *Educational Technology Research and Development*, 55(3), 223–252.
- Ho, W. Y. J. (2024). Digital multimodal composing pedagogy in a university writing course. *ELT Journal*, 78(3), 326–335. doi: 10.1093/elt/ccae011
- Hsiao, J.-C., Chen, S.-K., Chen, W., & Lin, S. S. J. (2022). Developing a plugged-in class observation protocol in high-school blended STEM classes: Student engagement, teacher behaviors and student-teacher interaction patterns. *Computers & Education*, 178, 104403. doi: 10.1016/j.compedu.2021.104403
- Jackson, H. M., Gulliver, A., Hasking, P., Leach, L., Batterham, P. J., Calear, A. L., & Farrer, L. M. (2024). Exploring student preferences for implementing a digital mental health intervention in a university setting: Qualitative study within a randomised controlled trial. *Digital Health*, 10. doi: 10.1177/20552076241277175
- Jewitt, C. and Kress, G. (2003). *Multimodal Literacy*. Bern: Peter Lang Publishing.
- Jiang, L. (George), & Hafner, C. (2025). Digital multimodal composing in L2 classrooms: A research agenda. *Language Teaching*, 58(4), 528–546. doi: 10.1017/S0261444824000107
- Kessler, M., & Marino, F. (2023). Digital multimodal composing in English language teaching. *ELT Journal*, 77(3), 370–376. doi: 10.1093/elt/ccac047
- Kim, K., & Kwon, K. (2023). Exploring the AI competencies of elementary school teachers in South Korea. *Computers and Education: Artificial Intelligence*, 4(September 2022), 100137. doi: 10.1016/j.caeai.2023.100137
- Kish, L. (1965). *Survey Sampling*. New York: New York, J. Wiley.
- Kong, S. C. (2014). Developing information literacy and critical thinking skills through domain knowledge learning in digital classrooms: An experience of practicing flipped classroom strategy. *Computers & Education*, 78, 160–173. doi: 10.1016/j.compedu.2014.05.009
- Kress, G. (2003). *Literacy in the New Media Age*. Routledge. doi: 10.4324/9780203299234
- Krueger, R. A., & Casey, M. A. (2015). *Focus Groups: A Practical Guide for Applied Research (5th Edition)*. Thousand Oaks: Sage.
- Lakens, D. (2022). Statistical Inferences and Effect Sizes: A Practical Introduction. In *Improving Your Statistical Inferences*. Retrieved from https://lakens.github.io/statistical_inferences/08-samplejustification
- Landis, J. R., & Koch, G. G. (1977). The Measurement of Observer Agreement for Categorical Data. *Biometrics*, 33(1), 159. doi: 10.2307/2529310
- Liang, J.-C., & Hwang, G.-J. (2023). A robot-based digital storytelling approach to enhancing EFL learners' multimodal storytelling ability and narrative engagement. *Computers & Education*, 201, 104827. doi: 10.1016/j.compedu.2023.104827
- Lim, F. V., Cope, B., & Kalantzis, M. (2022). A Metalanguage for Learning: Rebalancing the Cognitive with the Socio-Material. *Frontiers in Communication*, 7(February), 1–15. doi: 10.3389/fcomm.2022.830613
- Lim, F. V., & Unsworth, L. (2023). Multimodal composing in the English Classroom: recontextualising the curriculum to learning. *English in Education*, 57(2), 102–119.
- Lim, J., & Kessler, M. (2024). Multimodal composing and second language acquisition. *Language Teaching*, 57(2), 183–202. doi: 10.1017/S0261444823000125
- Linder, R., & Falk-Ross, F. (2024). Multimodal resources and approaches for teaching young adolescents: A review of the literature. *Education Sciences*, 14(9), 1010. doi: 10.3390/educsci14091010
- Liu, T., Gomez, G., & Shipman, F. M. (2025). Investigating children's multimodal enactment in digitally augmented tabletop storytelling. *International Journal of Child-Computer Interaction*, 46(August), 100770. doi: 10.1016/j.ijcci.2025.100770

- Maghsoudi, N., Golshan, M., & Naeimi, A. (2022). Integrating digital multimodal composition into EFL writing instruction. *Journal of Language and Education*, 8(1), 84–99. doi: 10.17323/jle.2022.12021
- Mayer, R. E. (2024). The Past, Present, and Future of the Cognitive Theory of Multimedia Learning. *Educational Psychology Review*, 36(1), 1–25. doi: 10.1007/s10648-023-09842-1
- Midgett, E., Stewart, O. G., & August, I. (2025). Understanding multimodal assessment practices in higher education to improve equity. *Education Sciences*, 15(11). doi: 10.3390/educsci15111523
- Mills, K. A., & Unsworth, L. (2017). Multimodal Literacy. In *Oxford Research Encyclopedia of Education*. Oxford University Press. doi: 10.1093/acrefore/9780190264093.013.232
- Mohammadi Zenouzagh, Z., Admiraal, W., & Saab, N. (2025). Empowering student engagement: the dynamics of learner traits in digital feedback environments. *Journal of Computing in Higher Education*. doi: 10.1007/s12528-025-09459-z
- Morgan, D. L. (1996). Focus Group. *Annu. Rev. Sociol.*, 22:129–52.
- Muzakkir, R., Saptarina, E., & Amelia, T. (2022). An analysis of students' difficulties in reading comprehension through online learning. *Linguistics: Journal Of Linguistics and Language Teaching*, 8(2), 258. doi: 10.29300/ling.v8i2.8322
- Nafidah, Novita, M., & Furqan, M. (2025). Impact of digital learning tools on student engagement and achievement. *Proceeding of International Conference on Education, Society and Humanity*, 4(1), 1436–1445. doi: 10.56778/jdlde.v4i1.511
- Najeeb, A., & Jameel, H. (2024). Difficulties face by students of BS English in Literary Analysis. *Pakistan Languages and Humanities Review*, 8(III). doi: 10.47205/plhr.2024(8-iii)09
- Nawi, M., Syafiq, N., Nor, M., & Amani, N. A. (2023). The challenges in the teaching of English literature. *Journey: Journal of English Language and Pedagogy*, 6(1), 130–147. doi: 10.33503/journey.v6i1.2640
- Nurjain, A., Nurjain, L. R., Fajriah, Y. N., Nurjain, A. K., & Firdaus, H. A. (2026). Integrating Generative Artificial Intelligence (AI)-Based Multimodal learning in education to enhance literacy aligned with Sustainable Development Goals (SDGs). *ASEAN Journal of Educational Research and Technology*, 5(1), 71–88. Retrieved from <https://ejournal.bumipublikasinusantara.id/index.php/ajert/article/view/765>
- Papen, U., & Peach, E. (2021). Picture books and critical literacy: Using multimodal interaction analysis to examine children's engagements with a picture book about war and child refugees. *The Australian Journal of Language and Literacy*, 44(1), 61–74. doi: 10.1007/BF03652065
- Puspitasari, H. R., Widiarti, N., & Subali, B. (2025). Digital storytelling for enjoyable and effective learning in the technological era (2020–2025). *Pedagogia: Jurnal Pendidikan*, 14(2), 161–173. doi: 10.21070/pedagogia.v14i2.1905
- Rahmanu, I. W. E. D., & Molnár, G. (2024). Multimodal immersion in English language learning in higher education: A systematic review. *Heliyon*, 10(19), e38357. doi: 10.1016/j.heliyon.2024.e38357
- Rençber, G. N., & Genç, G. (2024). Developing an achievement test on length measurement. *Education Mind*. doi: 10.58583/EM.3.3.8
- Rezigalla, A. A., Eleragi, A. M. E. S. A., Elhussein, A. B., Alfaifi, J., ALGhamdi, M. A., Al Ameer, A. Y., ... Adam, M. I. E. (2024). Item analysis: the impact of distractor efficiency on the difficulty index and discrimination power of multiple-choice items. *BMC Medical Education*, 24(1), 1–7. doi: 10.1186/s12909-024-05433-y
- Ryoo, J. H., Park, S., Suh, H., Choi, J., & Kwon, J. (2022). Development of a new measure of cognitive ability using automatic item generation and its psychometric properties. *Sage Open*, 12(2). doi: 10.1177/21582440221095016
- Sabirova, V., Azimova, Z., Asanova, D., Baytikova, B., & Miskichekova, Z. (2025). The impact of migration on the development of linguistic and literary competencies: Methodological approaches in philological education. *Social Sciences and Humanities Open*, 12(November), 102218. doi: 10.1016/j.ssaho.2025.102218
- Sahiruddin, Sudarwati, E., Lestari, I. P., & Daniswara, R. A. (2025). Digital multimodal literacy model of English language education in Indonesian higher education. *Journal of English Education and Linguistics Studies*, 2(2), 615–640. doi: <https://doi.org/10.30762/jeels.v12i2.2612>

- Sajjad, M., Iltaf, S., & Khan, R. A. (2020). Nonfunctional distractor analysis: An indicator for quality of Multiple choice questions. *Pakistan Journal of Medical Sciences*, 36(5). doi: 10.12669/pjms.36.5.2439
- Sakanti, D. P., & Widodo, P. (2020). Movie watching as visual literacy attainment. *LingTera*, 7(2), 149–156. doi: 10.21831/lt.v7i2.22813
- Shen, X., Shen, X., & Peng, J.-E. (2023). Digital storytelling: A literacy-building tool to promote willingness to communicate in a second language. *RELC Journal*, 55(3), 819–826. doi: 10.1177/00336882231157
- Silva, F. M. da, Siqueira, C. N., Neto, A. T. S., Castro, T. S. de P., & Olliveira, L. de. (2024). Multimodal literacy in the 21st century: Emerging perspectives and contemporary challenges. *RCMOS – Multidisciplinary Scientific Journal O Saber*, 1(1).
- Souza, R. R. de, Toebe, M., Mello, A. C., & Bittencourt, K. C. (2022). Sample size and shapiro-wilk test: An analysis for soybean grain yield. *SSRN Electronic Journal*. doi: 10.2139/ssrn.4082791
- Susilowaty, S., Setyarini, S., & Gustine, G. G. (2025). Crafting voices, shaping thought: integrating scaffolding strategies to foster critical literacy in selected young EFL writers. *REiLA: Journal of Research and Innovation in Language*, 7(2), 124–138. doi: 10.31849/reila.v7i2.26377
- Uhm, T., & Yi, S. (2023). A comparison of normality testing methods by empirical power and distribution of P -values. *Communications in Statistics - Simulation and Computation*, 52(9), 4445–4458. doi: 10.1080/03610918.2021.1963450
- Volker Pietsch. (2024). Into the multiverse – students commenting on audio-visual simultaneity. *L1-Educational Studies in Language and Literature*, 24(3), 1–29. doi: 10.21248/l1esll.2024.24.3.628
- Walldén, R., & Nygård Larsson, P. (2022). Joining the adventures of Sally Jones – Discursive strategies for providing access to literary language in a linguistically diverse classroom. *Linguistics and Education*, 72. doi: 10.1016/j.linged.2022.101121
- Walldén, R., & Winlund, A. (2025). Vocational Swedish for adult learners with limited educational background: rethinking literacy-based instruction. *System*, 135(August). doi: 10.1016/j.system.2025.103858
- Walsh, M. (2010). Multimodal literacy: What does it mean for classroom practice? *The Australian Journal of Language and Literacy*, 33(3), 211–239. doi: 10.1007/BF03651836
- Wigati, E. N., Lidyasari, A. T., & Adi, B. S. (2025). Leveraging digital learning tools to boost student motivation: A study using univariate and bivariate analysis on enhancing engagement in elementary science and social studies. *Mimbar Sekolah Dasar*, 12(1), 49–62. doi: 10.53400/mimbar-sd.v12i1.81864
- Wulandari, M., & Aslam, A. (2022). Hubungan antara literasi digital dengan hasil belajar siswa kelas sekolah dasar. *Jurnal Basicedu*, 6(4), 5890–5897. doi: 10.31004/basicedu.v6i4.3152
- Ylipulli, J., Pouke, M., Ehrenberg, N., & Keinonen, T. (2023). Public libraries as a partner in digital innovation project: Designing a virtual reality experience to support digital literacy. *Future Generation Computer Systems*, 149, 594–605. doi: 10.1016/j.future.2023.08.001
- Yu, S., Zhang, E. Di, & Liu, C. (2024). Research into practice: Digital multimodal composition in second language writing. *Language Teaching*, 1–17. doi: 10.1017/S0261444824000375
- Yuliana, E., Nirmala, S. D., & Ardiasih, L. S. (2023). Pengaruh literasi digital guru dan lingkungan belajar terhadap hasil belajar siswa sekolah dasar. *Jurnal Basicedu*, 7(1), 28–37. doi: 10.31004/basicedu.v7i1.4196

Appendix A: Pre-Post Test Instrument

A.1 Pre-Test Instrument

Instruction:

Read the text below carefully. Answer the following multiple-choice questions based on the story entitled "Putri Tanjuk Malam" (The Princess of the Evening Glow). Choose the most appropriate answer from options A, B, C, or D.

The Story of Putri Tanjuk Malam

This folktale originates from Telak Village, on the northern coast of Bangka Island. The story is not merely about a girl's beauty but also about honesty, tradition, and the consequences of poor choices. Through dialogue and narration, this story is revived as an oral literary heritage rich in local wisdom.

In ancient times, in an old settlement called Telak Village, located on the northern coast of Bangka Island, lived a peaceful and modest Malay community. They subsisted on farming and gardening, inheriting values of tradition, togetherness, and politeness that formed the breath of their daily lives. The village stood between expanses of pepper gardens and dense surrounding forests, holding many stories and ancestral beliefs.

In Telak Village, there lived a farming couple known for their kindness and respected by the community. They had a daughter of extraordinary beauty. Her beauty was not ordinary—it was said that when night fell, her face emitted a soft glow like moonlight. Because of this unique quality, the villagers called her "Putri Tanjuk Malam" (Princess of the Evening Glow).

"My daughter, you grow more radiant each night. What makes your face shine like the full moon?"

"Mother, perhaps it is only the lamp light... or perhaps it is because of your love and Father's love for me."

"There is no one in this village who does not know you. Even word of you has reached across the sea."

The nickname "Putri Tanjuk Malam" became legendary, spreading beyond Bangka Island. Days passed peacefully in Telak Village until one day, a wealthy merchant from across the sea arrived. Some said he came from Palembang, but most believed he came from Lingga Island, a territory of the Riau-Lingga Sultanate not far from Bangka. The merchant was known to be honorable, well-spoken, and possessed vast wealth.

His purpose for coming to Telak Village remained a question. Some said he came to trade, but many

believed he came specifically to propose to Putri Tanjuk Malam, for news of her beauty had reached the ears of the nobles in Lingga.

Upon arriving in Telak Village, the merchant met the princess's parents and stated his intention to marry the girl and take her to live on Lingga Island. However, the proposal was not welcomed with joy. The parents, though respectful of guests and maintaining Malay courtesy, politely declined. They did not want their only daughter to leave her homeland, especially for a distant and unfamiliar place.

"Sir, allow me to propose to your daughter, whose beauty has spread even to the palace. I promise to make her a queen in my land."

"Honorable Merchant, we are honored to receive a guest as noble as you. However... our daughter does not wish to be far from her ancestral land."

The princess peeked from behind the door, weeping softly.

"Mother... I do not want to leave. But I am afraid that refusing will hurt him."

"Be calm, my child. We will find a way that does not disgrace our customs and dignity."

Putri Tanjuk Malam shared the same sentiment. She was unwilling to leave her parents. However, bound by tradition and politeness, they did not want to refuse directly. So they devised a plan—a fake death drama to cancel the proposal without wounding anyone's pride.

When the merchant returned to Lingga to prepare for the wedding, the princess's family began digging a fake grave beside their house. Their plan was to bury the princess alive temporarily with the help of a bamboo tube as an air passage. They hoped that when the merchant returned and saw the princess's grave, he would believe the girl had died and abandon his intention.

Several days later, the merchant returned to Telak Village. He came with a large entourage, bringing various gifts, including bars of pure gold intended as a dowry. But what he found was not a wedding celebration but an atmosphere of mourning. The princess's parents wept and informed him that their daughter had died suddenly.

The merchant was stunned. He rushed to the princess's grave not far from the house. There, with a broken heart, he knelt and spoke softly,

"Alas, my beloved... you have departed."

As the word “departed” was spoken, lightning struck across the previously clear sky. Nature seemed to convey a message beyond human understanding. Inside the narrow hole where the princess hid, the air suddenly became heavy. Putri Tanjuk Malam struggled to breathe, and without anyone knowing, she breathed her last breath.

Unaware of what had happened beneath the earth, the merchant then placed the gold bars he had brought as a tombstone on the princess's grave. With a shattered heart, he left Telak Village, carrying an incurable wound.

Sometime later, the princess's parents, relieved that their plan had succeeded, intended to dig up the fake grave. However, how shocked they were to find the princess's body blue and stiff. The girl they loved had truly died. Weeping and sorrow broke out, accompanying deep regret for having toyed with fate.

They reburied her properly. Since then, it is said that on moonlit nights, the princess's grave emits a soft glow from beneath the earth—the light of her eternal beauty, a sign of regret, and a blessing that never fades.

The grave later became known to the community as the Keramat (Sacred Site) of Putri Tanjuk Malam, believed to be a sacred and mysterious place. Its location is now hidden deep in the forest of Telak Village, and only a few elders have ever known its whereabouts—one of them an old man who saw it as a child, but who no longer knows its exact location as the forest has changed.

Questions:

Choose the most appropriate answer from options A, B, C, or D.

1. What was the main purpose of the merchant's arrival in Telak Village?

- a. To open a new trading business
- b. To invite the villagers to sail to Lingga Island
- c. To propose to Putri Tanjuk Malam for marriage
- d. To offer modern agricultural assistance

2. What caused the death of Putri Tanjuk Malam?

- a. She was poisoned by her parents
- b. She died of profound grief
- c. She actually died from oxygen deprivation in the fake grave
- d. She was killed by the angry merchant

3. What is the main theme of the story “Putri Tanjuk Malam”?

- a. Rejection of arranged marriage by noble families
- b. The conflict between tradition and personal desire leading to regret
- c. True love fought for until the end of life
- d. Competition between villagers and foreign nobles

4. Which sequence of events is correct in the story?

- a. Proposal – Rejection – Princess dies – Merchant returns home
- b. Princess dies – Merchant arrives – Proposal rejected – Plan is made
- c. Princess is known for beauty – Proposal – Fake grave plan – Princess truly dies
- d. Merchant marries Princess – Princess dies – Parents regret

5. What is the main setting of the story?

- a. The palace of the Riau-Lingga Sultanate
- b. Telak Village on the northern coast of Bangka Island
- c. Lingga Island and its surrounding seas
- d. A noble's house in Palembang

6. Who is the main character in the story?

- a. The princess's parents
- b. The merchant from Lingga Island
- c. Putri Tanjuk Malam
- d. The old tomb keeper

7. What is the main message that can be taken from the story?

- a. Beauty can bring disaster
- b. A small lie can save one's honor
- c. No one can oppose God's will
- d. Leaving one's hometown is wrong

8. What moral value can be learned from the story?

- a. A polite refusal always brings good results
- b. Honesty and courage in facing reality are better than lies
- c. Following parents' wishes is an absolute obligation
- d. Rejecting someone's proposal will bring a curse

9. What life value is reflected in the story?

- a. Parents must fulfill all their children's desires for their happiness
- b. Honesty and courage in making decisions are better than deceiving to maintain tradition
- c. The princess should have accepted the proposal from the nobleman
- d. Living in a village is better than moving to another land

10. What is the most appropriate conclusion from the story?

- a. Proposals from nobles must always be accepted for the sake of honor
- b. Well-prepared plans will always succeed
- c. Humans cannot control fate, and lies can bring disaster
- d. It is better to leave one's homeland for personal happiness

A.2 Post-Test Instrument

Instruction:

Watch the animated video version of the story "Putri Tanjuk Malam" available at the link below. Then answer the following multiple-choice questions based on the video. Choose the most appropriate answer from options A, B, C, or D.

Video Link:

https://drive.google.com/file/d/1Dm93jlmYADMfetuOO6t53kjc_r_sNI7n9/view?usp=drive_link

Questions:

1. What was the main purpose of the merchant's arrival in Telak Village?

- a. To open a new trading business
- b. To invite the villagers to sail to Lingga Island
- c. To propose to Putri Tanjuk Malam for marriage
- d. To offer modern agricultural assistance

2. What caused the death of Putri Tanjuk Malam?

- a. She was poisoned by her parents
- b. She died of profound grief
- c. She actually died from oxygen deprivation in the fake grave
- d. She was killed by the angry merchant

3. What is the main theme of the story "Putri Tanjuk Malam"?

- a. Rejection of arranged marriage by noble families
- b. The conflict between tradition and personal desire leading to regret
- c. True love fought for until the end of life
- d. Competition between villagers and foreign nobles

4. Which sequence of events is correct in the story?

- a. Proposal – Rejection – Princess dies – Merchant returns home
- b. Princess dies – Merchant arrives – Proposal rejected – Plan is made
- c. Princess is known for beauty – Proposal – Fake grave plan – Princess truly dies
- d. Merchant marries Princess – Princess dies – Parents regret

5. What is the main setting of the story?

- a. The palace of the Riau-Lingga Sultanate
- b. Telak Village on the northern coast of Bangka Island
- c. Lingga Island and its surrounding seas
- d. A noble's house in Palembang

6. Who is the main character in the story?

- a. The princess's parents
- b. The merchant from Lingga Island
- c. Putri Tanjuk Malam
- d. The old tomb keeper

7. What is the main message that can be taken from the story?

- a. Beauty can bring disaster
- b. A small lie can save one's honor
- c. No one can oppose God's will
- d. Leaving one's hometown is wrong

8. What moral value can be learned from the story?

- a. A polite refusal always brings good results
- b. Honesty and courage in facing reality are better than lies
- c. Following parents' wishes is an absolute obligation
- d. Rejecting someone's proposal will bring a curse

9. What life value is reflected in the story?

- a. Parents must fulfill all their children's desires for their happiness
- b. Honesty and courage in making decisions are better than deceiving to maintain tradition
- c. The princess should have accepted the proposal from the nobleman
- d. Living in a village is better than moving to another land

10. What is the most appropriate conclusion from the story?

- a. Proposals from nobles must always be accepted for the sake of honor
- b. Well-prepared plans will always succeed
- c. Humans cannot control fate, and lies can bring disaster
- d. It is better to leave one's homeland for personal happiness