iejee

September 2023, Volume 16, Issue 1

www.iejee.com

international electronic journal of elementary education

KURA group Education & Publishing

INTERNATIONAL ELECTRONIC JOURNAL OF ELEMENTARY EDUCATION

ISSN: 1307-9298

www.iejee.com
iejee@iejee.com

KURA group
Education & Publishing

iejee

International Electronic Journal of Elementary Education
September 2023, Volume 16, Issue 1

Table of Contents

ISSN: 1307-9298

iejee

International Electronic Journal of Elementary Education
September 2023, Volume 16, Issue 1, 1-16

Self-Regulatory Metacognitive Skill Use in Elementary Students During Computer and Paper Reading Assignments: A Qualitative Study

Katerina Sergi[a,*], Anastasia Elder[b], Tianlan Wei[c], Kristin Javorsky[d], Jianzhong Xu[e]

Received : 6 June 2023
Revised : 3 September 2023
Accepted : 29 September 2023
DOI : 10.26822/iejee.2023.310

[a,*] **Corresponding Author:** Katerina Sergi, Mississippi State University, USA.
E-mail: ks1894@msstate.edu
ORCID: https://orcid.org/0000-0003-3777-0327

[b] Anastasia Elder, Mississippi State University, USA.
E-mail: aelder@honors.msstate.edu
ORCID: https://orcid.org/0000-0002-9274-1572

[c] Tianlan Wei, Mississippi State University, USA.
E-mail: tw1518@msstate.edu
ORCID: https://orcid.org/0000-0002-8270-5287

[d] Kristin Javorsky, Mississippi State University, USA.
E-mail: kj911@msstate.edu
ORCID: https://orcid.org/0000-0001-6328-1442

[e] Jianzhong Xu, Mississippi State University, USA.
E-mail: jx18@colled.msstate.edu
ORCID: https://orcid.org/0000-0002-0269-4590

Abstract

Self-regulated learning (SRL) and metacognitive processes are important in education because they contribute to effective learning and improved academic performance. Metacognitive SRL may be facilitated by the implementation of computer technology. This qualitative study examined the presence and use of metacognitive SLR processes among elementary school students as they completed computer- and paper-based reading assignments. Students in two after-school programs were recruited from a public school district in a southeastern region of the United States (U.S.). The participants consisted of 52 elementary students in Grades 2-5. Students participated in two, counterbalanced, conditions that involved computer- and paper-based reading assignments. Observations and semi-structured interviews were conducted. The results indicated that students were more likely to apply metacognitive SRL skills when reading on paper than reading on a computer. Overall, students showed signs of *planning* more in the paper than in the computer condition but student behaviors and responses differed between grades. *Monitoring* practices appeared in both the computer- and the paper-based reading assignment, with monitoring connected with background knowledge in Grades 2 and 3 but reading content in Grades 4 and 5. *Control* processes such as retrying and representing graphically were more common in the computer- than in the paper-based reading across all grades. Students used their score in a reading assignment as an *evaluation* tool to assess performance in the computer- and paper-based reading condition. These findings suggest that the utilization of prior information, integration of multimedia and verbal signals, and comfort level with the reading medium all influenced students' SRL decision-making.

Keywords:
Self-Regulated Learning, Metacognition, Online Reading, Elementary Grades

Introduction

Investments in education that are tied to computer technology in classrooms have increased recently (Gray & Lewis, 2021; Smith, 2016). Due to these initiatives, one computer per student is now virtually universally used (The White House, 2021; U.S. Department of Education, 2016; Zheng et al., 2016). The use of computers in the classroom enhances teaching and learning. In the domain of reading, computer technology is widely used. Studies have shown that digital reading can contribute to the development of language skills (Mayer, 2003; Xu et al., 2021), foster active learning (Dalgarno & Lee, 2010; Fiorella & Mayer, 2016; Robinson, 2016), and increase interest (Schiavo et al., 2021). Yet, such challenges as adaptive learning, prolonged engagement, task processing stemmed from digital modalities exist (Greenhow et al., 2022; Máñez et al., 2022). Also, variation in reading comprehension between grades and digital and print media (Clinton, 2019; Furenes et al., 2021), differences in achievement scores (Gulek & Demirtas, 2005; Mora et al., 2018), but also in higher-order thinking have been reported (Combrinck & Mtsatse, 2019; Earle et al., 2020). Despite the growing popularity of computer technology in education, there is limited evidence that primary school students use higher-order cognitive skills for text comprehension in their typical classroom and computer lab environments. Incorporating self-reflections and learning strategies into computer-based classrooms implies that learners would have to assess the thought process behind their conclusions; this process is known as metacognitive SRL. The current qualitative study attempts to fill this gap by providing naturalistic accounts of students' completing computer- and paper-based reading assignments. The findings may provide insights into ways elementary students apply SRL metacognitive processes and may have implications for instruction and educational technology.

The Interplay of Self-Regulated Learning and Computer Use

Planning, monitoring, control strategies, and evaluation are all part of the SRL metacognitive processes (Bandura, 1991; Pintrich & de Groot, 1990; Schunk, 2008). Planning is the process of thinking strategically and taking proactive steps to achieve a learning objective (Corno, 1994; Greene et al., 2010). Monitoring is the cognitive response that propels learning objectives while controlling thought and behavior (Brown, 1977; Pintrich, 1999). Metacognitive SRL entails active control and corrective actions, as well as learning tactics like concept organization, close reading, reading aloud, and retrying (Flavell, 1979). Evaluation involves the performance mechanism by which a learner assesses if the desired goals have been met (Koriat, 2012; Manlove et al., 2007; Zimmerman, 1989).

Examining SRL metacognitive indicators can help us understand the ways scaffolding, modeling, feedback, and interactivity assist students while engaging in computer tasks (Bannert et al., 2015; Serrano et al., 2018; Sha et al., 2012). Researchers have suggested that the integration of computer technology in education is important as computers can encourage the transfer of prior knowledge (Bulu & Pedersen, 2012; Muis et al., 2015; Price & Oliver, 2007). Notable effects using computers in education have been reported in relation to memory gains and attention (Chevalère et al., 2021; Schacter & Szpunar, 2015); use of higher-order executive skills for goal attainment (Al-Jarrah et al., 2018); problem-solving development in middle-school students (Muis et al., 2016; Postholm, 2011), and in high-school students metacognitive SRL during online courses and in math-learning software (Dresel & Haugwitz, 2008; Lin et al., 2017). Conversely, digital reading may produce taxed working memory and cognitive load (Paans et al., 2018; Pratt & Martin, 2017), miscalibration between perceived and actual accuracy in performance tests (Pilegard & Fiorella, 2016), or boredom and frustration that may cause distraction in goal achievement (Artino, 2009). In summary, metacognitive SRL develops during the middle-school years and involves the ability to make autonomous learning choices, to adapt goals to new circumstances, and modify responses after receiving positive or negative feedback.

Self-Regulated Learning and Computer- versus Paper-Based Reading

In the present study, we attempt to shed light on students's reading comprehension. The process of learning entails coming to conclusions through self-reflection and self-regulatory strategies (Groß, 2021); these processes are also essential in reading comprehension (Earle et al., 2020; Qi, 2021). The most appropriate way for students to extract meaning during reading is through conscious and controlled use of reading strategies, which requires a degree of metacognitive skills (Koutsouraki, 2020; Pressley & Afflerbach, 1995). Notably, reading comprehension is associated with increased SRL practices (C.-M. Chen et al., 2019; Q.-S. Chen, 2009). Scaffolded supports have been positively associated with SRL processes and metacognitive strategies in computer-based educational environments (Serrano et al., 2018; Vidal-Abarca et al., 2010). Past research has demonstrated that scaffolding strategies in online environments can boost the metacognitive skills of sixth graders, and these skills depend upon the use of prior knowledge in upper-grade students (Bulu & Pedersen, 2012; Roussel, 2011). But, in a meta-analysis of 54 studies between 2000-2017, Delgado et al. (2018) compared print and digital reading for children and adults and pinpointed a digital inferiority of the computer medium in fostering reading comprehension and

learning tasks. Furthermore, Clinton (2019) conducted a meta-analysis on 33 studies that examined paper versus screen reading for children and adults during a decade, 2008-2018. Their findings suggested that paper reading constituted a more efficient way to comprehend the material and improve test performance than screen reading. Furenes et al. (2021) also meta-analyzed findings from 39 studies that focused on paper and digital reading in children ages 1 to 8 years old. Their findings corroborated previous results that showed lower comprehension rates of digital than paper reading. However, the authors professed that digital reading that contained visual and story vocabulary cues outperformed paper reading. And, Latini and Bråten (2022) researched a sample of 116 Norwegian undergraduate students in relation to reading informational texts on a tablet versus paper. The results did not favor the hypothesis of paper offering comprehension advantages. In fact, there were no differences between the two media in terms of metacognitive cognitive and behavioral activities during text processing. Therefore, examining metacognitive SRL practices of elementary students in computer-based versus paper-and-pen reading tasks can advance our knowledge-base regarding differences in development and impact of the reading medium.

From a methodological perspective, observations and in-depth interviews can produce deep explorations of SRL metacognitive practices in computer-supported learning (Ferreira et al., 2017; Postholm, 2011; Robson, 2016). Past qualitative approaches to examining computer-based SRL include: (a) discourse analysis of virtual learning interactive communities where a reconciliation of individualized with collaborative learning was enacted (Delfino, et al., 2008); (b) triangulation of interviews from elementary and middle school students, teachers, and administrators which revealed positive correlations between personalized computer learning and persistence in completing reading assignments (Underwood & Banyard, 2008); (c) case study for primary school students where video-engaging recall produced fewer monitoring activities in reading (Pratt & Martin, 2017); (d) transcriptions of students' speaking aloud utterances demonstrating that student engaging in a hypermedia-learning environment contributed to deep-strategy use (Deekens et al., 2018). The current study extends previous qualitative research by incorporating natural accounts of elementary-aged students as they complete computer and paper-pen reading assignments.

Research Questions

This qualitative study had two research questions: (1) "Do elementary students demonstrate the use of SRL metacognitive processes during computer-based and paper-pencil reading tasks?"; and (2) "Are there apparent differences in the use of SRL metacognitive processes between the primary and upper elementary grades?"

Methods

Epistemological Paradigm

We applied the epistemology of constructivism to shed light and meaning into the reality of computer technology use in elementary education (Krauss, 2005; Guba & Lincoln, 1994). We conducted observations and semi-structured interviews to (a) record actions and utterances in a naturalistic educational setting; (b) gain a deeper level of detail by establishing rapport with each participant; and (c) eliminate dominating or distracting voices, thus allowing a variety of perspectives and ideas to surface. Observations and semi-structured interviews were conducted concurrently (Creswell & Plano Clark, 2010; McCrudden et al., 2019). Research assistance from trained undergraduate students contributed to the triangulation of methods and data and guarded against researcher bias (Tashakkori & Teddlie, 1998).

Recruitment

We recruited elementary students in Grades 2-5 from two after-school programs in a southeastern U.S. school district. The first program was fee-based, served students from pre-K to grade 5 operated 5 days a week during the school year, and provided recreational and enrichment activities. The second after-school program was funded through the 21st Century Community Learning Centers program, served students from Kindergarten to grade 5, operated 3 days a week, and emphasized academic tutoring. The after-school programs gave us access to elementary students during non-school hours, thus affording us uninterruption of regular instruction. The study was approved by the school district and the university's IRB. We invited 156 parents, and 69 agreed to have their children participate (42% return rate). Out of 69 consented students, data collection produced 52 accounts and concluded in spring 2020 when the COVID-19 lockdown forced the suspension of all school activities.

Sample

The 52 students were from 32 different classrooms with different teachers. The final sample size was acceptable for achieving saturation for the underlined theoretical associations (Baker & Edwards, 2012; Saunders et al., 2018). Saturation was achieved when codes, themes, and theoretical components were repeated, and no new elements of relevance were revealed through additional data (Glaser & Strauss, 1967; Saunders et al., 2018). Participants' age ranged

from 7 to 11 years-old (M = 9.06, SD = 1.21; 62% boys), and 29% were African-American, 6% Asian-American, and 65% European-American. As an indicator of socioeconomic status (SES), we used participation in a free or reduced meal program. Among participants, 21% received free lunch, 4% received reduced lunch, and 75% paid for their lunch (Table 1).

Measures

To conduct observations and semi-structured interviews, we employed items from the Junior Metacognitive Awareness Inventory (Jr. MAI; Sperling et al., 2002) and the original MAI (Schraw & Dennison, 1994). The subscale regulation of cognition was used for assessing the following dimensions: planning (3 items – "think before I choose," "think of several ways," and "make side notes", including annotating and pointing at words with cursor or pencil); monitoring (2 items – "ask how well I am doing," and "check before moving on"); control-learning strategies (3 items – "draw diagrams," "pay attention to cues," and "read out aloud"); evaluation (4 items – "go over unclear information," "reread," "ask myself if I learned," and "know how well I did"). Items were adapted to reflect engagement with computer- and paper-and-pencil-based reading assignments. During the observation phase, we recorded SRL metacognitive behaviors and questions from participants and notes and memos from researchers. During the interview phase, we asked participants about their SRL metacognitive behaviors and cognitive processes using prompts and follow-up questions.

Procedures

Each student attended two sessions: Condition 1 – reading assignment using computer; and Condition 2 – reading assignment using paper and pencil. Sessions were separated by 7-15 days. The order of conditions was counterbalanced and randomly assigned across students; in each grade, about half of students first completed the computer condition and half the paper-pencil condition. Each participant was observed and interviewed alone. In both conditions, students completed a reading assignment in a classroom lab. Observations and interviews were conducted by the principal researcher with support from four undergraduate assistants, one per session. Assistants were blinded to each other's results and to principal researcher's. Reading tasks in the computer condition were assigned by students' teachers in their respective grade using the i-Ready online educational program which aligns with Common Core Standards (Curriculum Associates LLC., 2019). Two distinctive features of the online version of i-Ready are narration and feedback. Reading tasks in the paper-and-pencil iReady condition were assigned by the principal researcher. Students completed an i-Ready worksheet for reading based on curriculum and standards guidelines by the state Department of Education and the local school district for exemplar units and lessons for the academic year. Reading topics in each condition ranged from myths, fiction, and poetry, to scientific facts, narrative and expository texts. During each session, participants were observed for SRL metacognitive practices, think-aloud utterances, and questions. Then, the researchers conducted interviews with students about their SRL metacognitive behaviors. Each session of observation and interview lasted on average 30 minutes. Items from Jr. MAI and MAI were read to participants and sometimes rephrased for clarification. To describe students' reflections on SRL metacognitive processes and gain insights on direct and retrospective cognitive and behavioral practices,

Table 1

Demographic Profile of All Participating Students in Grades 2-5 (n = 52)

Demographic Profile	2nd grade Count (%)	3rd grade Count (%)	4th grade Count (%)	5th grade Count (%)	Total Count
Sex					
Boys	6 (67%)	9 (64%)	12 (71%)	5 (42%)	32
Girls	3 (33%)	5 (36%)	5 (29%)	7 (58%)	20
Race					
African-American	3 (33%)	5 (36%)	2 (12%)	5 (42%)	15
Asian-American	0 (0%)	0 (0%)	2 (12%)	1 (8%)	3
European-American	6 (67%)	9 (64%)	13 (76%)	6 (50%)	34
Free or Reduced Lunch					
Free	2 (22%)	3 (21%)	1 (6%)	5 (42%)	11
Reduced	0 (0%)	1 (7%)	1 (6%)	0 (0%)	2
Paid	7 (78%)	10 (72%)	15 (88%)	7 (58%)	39
Age Breakdown					
7 years	5 (56%)				5
8 years	4 (44%)	10 (71%)			14
9 years		4 (29%)	9 (53%)		13
10 years			8 (47%)	5 (42%)	13
11 years				7 (58%)	7

we asked follow-up, open-ended questions (e.g., "Why do you say that?", or "Could you tell me more about that?"). We recorded behaviors and thinking processes related to a specific reading assignment given during the session, as well as to a reading assignment typically given at school. Memos from the principal researcher and the undergraduate students were also included in the data.

Data Analysis

Qualitative analysis was conducted using the NVivo 12 software (QSR International, Melbourne, Australia). The data analysis involved the following stages: (a) listening to audio recordings and transcribing them; (b) reading transcripts to highlight comments or phrases representative of participants' perspectives; (c) clustering of highlighted statements into summaries for generating domains of meaning; (d) classifying data sources by type (e.g. semi-structured interviews, observation notes, researcher memos); (e) identifying sections corresponding to general concepts or themes; and (f) aligning original codes with key SRL metacognitive processes. Once all statements were coded, four major classifications (themes) emerged based on the theoretical framework of SRL and metacognition; these were planning, monitoring, control, and evaluation. Analyzing these data allowed for reflections on emerged themes and comparisons between the two conditions; these themes included sounding out, and distractions. The data were closely examined to create lists of codes. New codes were constantly compared to earlier codes, and final codes were integrated into the theoretical constructs used. As codes integrated into main themes, we checked for additional sub-codes and performed axial coding to further collapse the data (Miles & Huberman, 1994). In addition, we cross-referenced codes with all data sources (i.e., observations, interviews, and memos) to deeply explore the data. The primary analytic strategies used were thematic analysis and triangulation. The researchers reviewed coded sections and looked for emerging themes across data sources according to the constant comparative method (Corbin & Strauss, 1990; Glaser & Strauss, 1967). Then, the researchers checked for converging or conflicting findings by groups of participants and data collection method.

Trustworthiness and Authenticity

Triangulation of data (i.e., field observations, semi-structured interviews, and memos from multiple researchers) was employed to achieve transferability of conclusions and trustworthiness (Guba, 1981; Guba & Lincoln, 1994; Kornbluh, 2015). Credibility was ensured by instantaneous checking data accuracy as previously recommended (Shenton, 2004). To ensure dependability we examined the raw data in each condition and against the researchers' personal reflections in memos and in audit trails that evaluated the effectiveness of the qualitative inquiry and implementation. To ensure trustworthiness, we: (a) triangulated the sources; (b) employed multiple raters; (c) matched the data with pre-determined theoretical constructs and created additional codes for emerged themes; (d) engaged with the data through a prolonged analytic process and the writing of rich descriptions; and (e) applied member checking of behaviors observed in interviews. To ensure authenticity, we: (a) used original research objectives; (b) maintained a consistent demeanor when observing and interviewing students; (c) involved original and follow-up questions to account for evolving themes; (d) defined constructs with an operational codebook; (e) perfomed critical appraisal and interpretation of the data (Guba & Lincoln, 1994; Morrow, 2005; Yin, 2011).

Measures Against Researcher Bias

To protect against researcher bias, we treated students as active agents with their own voice (Howitt & Cramer, 2011). The principal researcher guided conversations using simple language and specific prompts. Moreover, students participated at their own free will in accordance with ethical research guidelines. Trustworthiness and credibility were safeguarded as described earlier. Finally, the researchers engaged in self-reflection and self-questioning using memos and personal comments to eliminate traces of implicit bias (Wilson, 1998).

Results

Next, we present our findings in each metacognitive SRL dimension, namely planning, monitoring, control, and evaluation processes.

Planning

In the computer condition, second and third grade students first read a passage in its entirety and looked for word definitions, then responded to questions. Students in Grades 3 and 4 pointed the cursor to words to direct eye-text movement. When asked, third graders replied that they intuitively used response elimination but they could not explain why. Statements from fourth grade students suggested a reliance on teacher instruction and an emphasis on specific vocabulary words as planning tactics. In general, students did not annotate on a separate piece of paper. Remarkably, third and fourth graders perceived annotations as "cheating" because of not using their "own brain" but an external aid. Fifth graders relied on computer characters for story narration or question citation, a tactic that resulted in responding without reading. However, fifth graders claimed that computer narration helped them comprehend word definitions and pronunciation and validate their own thinking process:

> *The text helps and the way that the computer character speak also helps me think about the way I say it [and] the way they say it. I kind of combine it together and I know how to say it. For the "Write" portion, I usually answer the question and then I know that since I've answered the question, let him [computer character] know what I'm talking about. And then I just use the rest of my own words.*

During the paper-based assignment, second graders finger- or pencil-pointed while reading, or numbered verses in poems. Third and fourth graders approached paper reading with a plan: some first checked all text boxes and passage in the worksheet, whereas others looked at all questions first and answered them second. In addition, third graders mentioned tangible elements for planning their reading in print (e.g., looking at title or identifying context clues, and then "circling the best response", or "eliminating at least half of those [responses] that might be wrong.") Fourth graders appeared influenced by the reading genre and paused to integrate previous learning before responding to questions. Unlike computer reading, fourth graders created annotations, especially for reading texts perceived as "hard." Finally, fifth grade students viewed reading passages in their entirety, pondered, and then started answering questions. They notably avoided annotations because they perceived them as "off-track" distracting activities.

Students across conditions and grades used a structural approach to reading and pointed either the computer cursor or pencil to focus and complete a reading assignment. However, we identified the following variations in planning between conditions: (a) fourth graders integrated prior knowledge for completion more in the paper than the computer condition; (b) third graders used more context clues in the paper than in computer condition; (c) third grade students applied the process of elimination in both conditions; (d) fourth graders made annotations in both conditions; and (e) fifth grade students listened to computer narration.

Monitoring

During the computer-based reading assignment, second grade students asked themselves questions and repeated difficult words to crystallize understanding. Another monitoring method was to ask the researchers questions about the writing segment of a reading assignment, or about the meaning or spelling of a word. Third graders used rechecking and rereading after receiving a bad grade in a test. Rereading prompted students to rely on computer character narration because it "helps to understand and get it right." Fifth graders, however, perceived computer character narration as confusing because it contradicted their prior knowledge. Those students talked to themselves to recall previously learned elements in the text and then continued reading to understand. Fourth and fifth graders commonly used computer feedback for monitoring and in revising responses:

> *"If I can't remember what I answered wrong, the program tells me when I answer a question wrong. When I don't understand it, I might read it a couple of times. And then, I might see if the computer will read it, just to clarify what it is."*

During the paper-based reading assignment, students either asked themselves or asked the researchers questions:

> *As a student read multiple choice options, he said "maybe" for the correct option. In the case of open-ended questions, the student repeated the phrase "which item supports the point" before he answered. The student tilted his head up and whispered, "what was I thinking about that?" After the student read options again, he posed a question by asking "so what was the question again?" [Researchers' observation notes for second grade students]*

Students focused on the meaning and pronunciation of challenging words while completing paper-based reading assignments. Students in Grades 2 and 3 verified responses upon teacher review. However, fourth graders said that "second-guessing" was a disadvantage and resulted in wrong responses "because a lot of times when I do that, I get it wrong when I [at first] had the right answer." Furthermore, most fourth and fifth graders were concerned with "how to do things" in completing open-ended questions, reading aloud, recording evidence from passages, or graphically representing meaning.

Overall, students demonstrated monitoring practices in both the computer- and paper-based reading assignments, i.e., asking questions, and rechecking answers. Monitoring was knowledge-based in Grades 2 and 3 meaning prior knowledge of content affected the process of monitoring. Conversely, monitoring was task-based in Grades 4 and 5, meaning familiarity with condition affected the process of monitoring. There were variations between conditions and grades in monitoring practices. Fourth and fifth grade students used computer-generated feedback to a greater extent than students in other grades. Third grade students used informational clues more in the paper than in the computer condition.

Sounding out

A sub-theme of monitoring documented more in the paper than in the computer condition, was "sounding out." According to students, "sounding out" meant piecing together a word from its syllables. In the computer-based assignment, "sounding out" occurred with computer-generated feedback. For second and third graders, "sounding words out" was a way to understand ambiguous reading elements and correctly complete the assignment. Fourth grade

students relied on the computer for pronunciation and meaning of unknown words: "if it's important, I'll click on [word] and [computer program] will have a speech thing. I'll see what it is and the definition of the word." However, fifth graders stated that when tried "sounding out" this strategy did not improve their comprehension.

Second, third, and fourth grade students retrospectively mentioned that, in the paper-based reading assignment, "sounding it out" helped them understand unknown or unclear words. Among second graders, "sounding out" words was not beneficial; however, for third grade students, sounding out helped "figure out what the answer is" and contributed to close reading strategies (e.g., rereading, reflecting). Fourth graders mentioned that examining the parts of a word aided in content comprehension of paper reading assignments: "I look at the beginning of the word and what that means, then I look at the end of the word, and then I figure out what the word means." Other fourth graders mentioned that "sounding out" was a multi-step process: "I'll cover up part of the words and say 'participation.' I'll cover up 'icipation' and I say, 'I know that word is 'part' and then I'll figure out the next part."

Control

Students demonstrated some control processes during computer-based reading, one of which was organizing information in charts and tables. Fourth and fifth grade students acknowledged that creating pictures or diagrams left mental traces that aided in remembering information. Yet, when asked, students across grades reported that tables/charts did not aid comprehension of computer reading assignments. Students in Grades 2 and 3 stated that information on sidebars and context clues assisted them with deciphering important information. Likewise, students in Grades 4 and 5 used information cues and "read it" buttons to review and understand computer-based reading passages. This helped students to "almost always get the questions right." Surprinsingly, computer read aloud tools assisted students with reading text.

During the paper-based reading assignments, students demonstrated some control processes. Table/chart creation was more challenging in the paper than the computer condition for early elementary students. Many of them did not understand how to do it and skipped parts of reading assignments. Fifth graders successfully created organizational charts but time restrictions prevented from completing them. In the paper condition, students used surface-level strategies, e.g., underlining or circling; this was mostly evident among second graders. Third graders tended to pause and ponder as they gradually completed the writing portion of paper reading assignments, whereas fourth graders read titles, bold words, and hint boxes. Response elimination (i.e., crossing out response choices and circling the correct or best response) was a prominent comprehension control strategy for fifth grade students. Reading aloud was hardly evident across students. Students in Grades 2 and 3 claimed that they did not want "to give away" their answers and attested that "reading in my head is more helpful" and this way "I would not interrupt the class." Other fourth graders seemed to associate reading aloud with an emotional state because "if I can't do it, it makes me more stressed out, and when I can read out loud, I can understand it while I have to process it in my head reading it." But with incorrect selections, voicing words silently and retrying were observed in third, fourth, and fifth graders.

Generally, students across grades were keener in retrying and organizing during computer than paper reading, but they performed close reading mostly in paper reading assignments. Reading aloud was barely documented in any of the two conditions, whereas the use of visual cues facilitated retrying in both conditions. During both the computer and paper conditions, visual elements seemed to mediate students' review and retry. Visual elements functioned as comprehension triggers to check learning. These features included checkmarks, smiley faces, and praise from animated characters in the computer condition, but bold words and informational text boxes in the paper condition. In both conditions, students in Grades 2, 3, and 4 stated that they had not been taught schematic representations of content. Most younger students could not recognize relationships, identify patterns, and establish connections in text. Nevertheless, computer embedded charts facilitated student learning because they made content easier to understand and remember. Yet, computer embedded charts discouraged students from creating charts on their own.

Distractions. A sub-theme of control that emerged was "distractions," and more so in computer than in paper reading. A form of distraction was eye regression – defined by Squire (2009) as the backward movement of the eye when reading. In our study, eye regression represented eye movement away from the text which seemed to interrupt SRL practices and independent learning. Eye regression was more evident in older than younger students. Second graders relied on computer character narration of text or questions and responses which contributed to eye regression. Fourth graders seemed to divert their attention from the computer screen and look elsewhere, or fidgeted with irrelevant objects (e.g., own glasses, headphone cord). Likewise, eye regression in fifth grade students resulted in performing no SRL practices and randomly responding without reading. Some third-grade students crossed hands or held their face in boredom in

response to lengthy passages and difficult vocabulary during the computer-based reading assignment. Students reported that completing computer reading assignments decreased concentration. However, older students perceived "distractions" as advantageous because they triggered them to seek assistance from teachers.

Evaluation

In computer-based reading, students used embedded features to evaluate their performance; however, there were small variations between grades. To check the accuracy of a response, students relied on interactive multimedia including "green highlighted text" or "confetti throws" for correct responses, or green "DONE" signs for completed responses. Furthermore, praise prompts such as "beautiful," "you got it," "good job," or "nice one" enabled second-grade students to evaluate their progress. Third graders evaluated their performance using their progress score: "because one time, I didn't check my answer, I did mostly move on to the question, but then I got a '71' because I didn't go back and now I'm starting to go back and see." Fourth and fifth graders used computer feedback for evaluation when reviewing responses. Other fourth graders clicked on different response options to hear them in a sentence and then proceed to selecting using "trial and error."

During paper-based reading, students evaluated their performance by returning to the text. Repetitive going back was evident in the writing portion of the assignment, especially for students in Grade 3 or higher. Third graders stated that new information necessitated to "look at words around it" for clarity. Fifth graders seemed to rely on memory skills because they checked back less often than students in earlier grades. There were no instant evaluation prompts in paper-based reading. Students across grades reported that they did not always know their progress on paper reading assignments, unless the teacher graded it instantly and marked it with "smiley faces."

Overall, students across grades used their progress score to assess their performance in the computer and paper condition. However, we recorded two major differences in evaluation between the two conditions. First, students reported that they received continuous feedback and an instantaneous score during computer-based reading, whereas teacher feedback and score were delayed during paper-based reading. A faster evaluation seemed to provide more opportunities for corrective actions in computer- than paper-based reading. The second difference was that students return to the text to review and evaluate their performance more often in the paper- than the computer-based reading assignment.

Summary of Qualitative Findings

Students across all grade levels demonstrated SRL metacognitive processes under both reading conditions, although primary grade students were more likely do so when reading on paper. Specifically, students showed more signs of planning in the paper than in the computer condition but student behaviors and responses differed between grades. Monitoring practices appeared in both the computer- and the paper-based reading assignment; monitoring relied on prior content knowledge in Grades 2 and 3 and reading medium in Grades 4 and 5. Control processes such as retrying and organizing were more common in the computer- than in the paper-based reading. Close reading habits appeared more in paper than in computer reading, whereas distractions influenced control strategies more in computer- than in paper-based reading. Students used their progress score for evaluation of performance in both conditions.

Figure 1 highlights the main ideas associated with SRL metacognitive dimensions in the computer- and paper-based reading conditions.

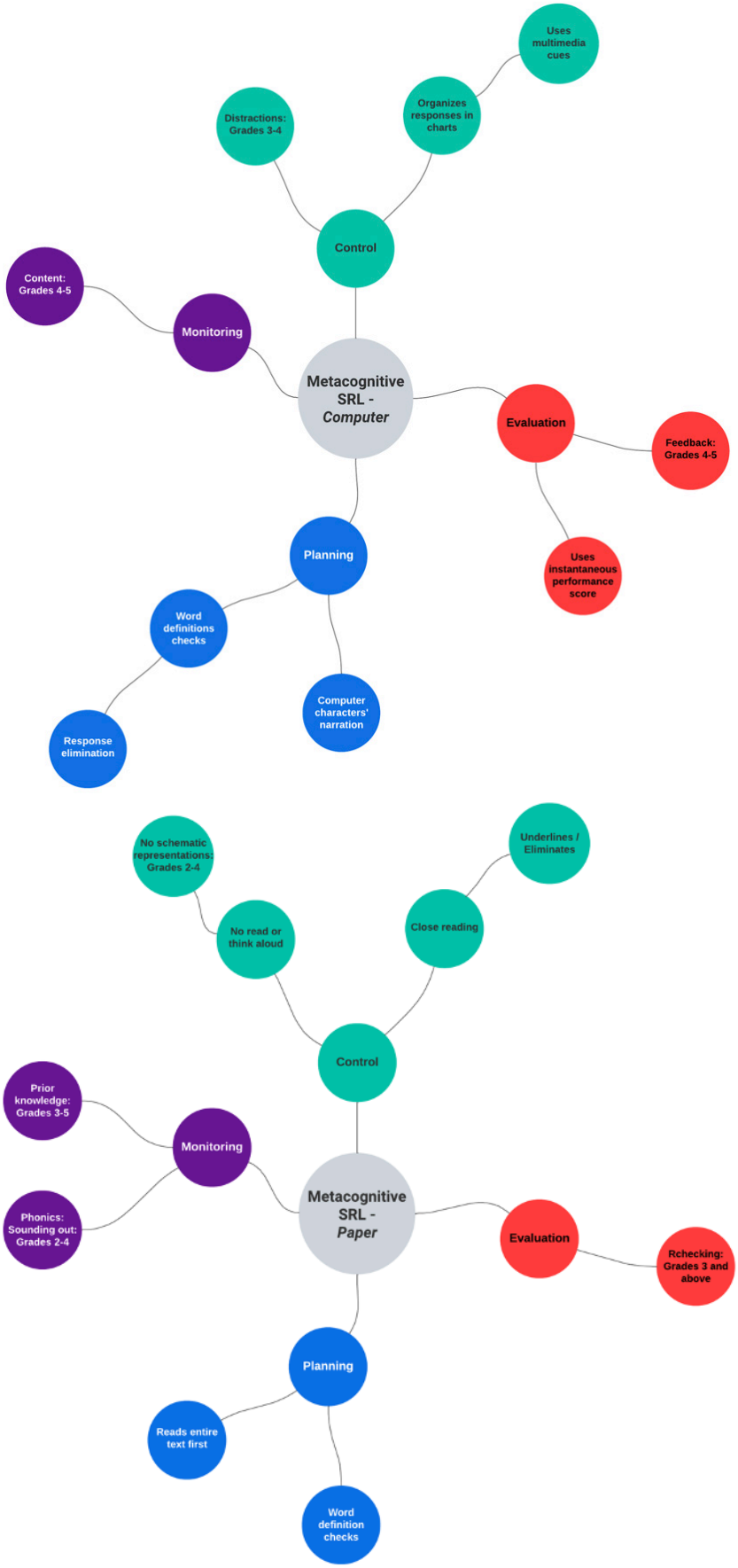
Discussion

The present qualitative study examined if elementary students in Grades 2-5 demonstrate SRL metacognitive processes during computer- and paper-based reading comprehension tasks and whether differences in SRL decision-making between conditions exist across grade levels. We documented that while students across all grade levels demonstrated SRL metacognitive processes under both conditions, strategy selection varied by grade level, and was overall more prevalent during paper-based reading.

The elementary students of the present sample regulated cognition by applying more types of planning and control processes in the paper than in the computer condition, and by demonstrating monitoring and evaluation processes in both conditions. Previous research has consistently shown the effects of planning to be present in older students (middle, high school, and undergraduate students) during both computer- and paper-based reading assignments (Follmer & Sperling, 2019; Manlove et al., 2007). Elementary students, however, have not been found to use planning processes when using electronic media (Muis et al., 2016), which may be attributed to differences in task and medium. Complementing this research, our results suggest that students apply planning to purposefully set learning goals and use strategies for attaining these goals. This finding is demonstrated more so in paper- than computer-based reading assignments because students may be more accustomed with organizing their learning and completing reading assignments in paper format (Greene et al., 2010; Kuisma & Nokelainen, 2018).

Figure 1

Main Findings of Metacognitive SRL Processes Between Computer and Paper Reading Assignments in Grades 2-5



Students in our study demonstrated specific judgements that lead to monitoring practices in both computer and paper reading such as going back to the text in accordance with previous research in high school students (Hardcastle et al., 2017). Also, the present students engaged in the control strategy of close reading more in the paper than in the computer condition likely because they seemed to understand the presented features better during paper- than computer-based reading, which is consistent with past findings in middle and high school students (Kim & Kim, 2013; Stoop et al., 2013). Furthermore, we saw that students paid attention to visual and context clues to monitor reading deficiencies and apply learning strategies. However, previous studies found more difficulty in the comprehension of text in screen/computer-based reading, which could also lead to more need for SRL strategies during monitoring and control (Clinton, 2019; Delgado et al., 2018; Furenes et al., 2021; Latini & Bråten, 2022). Similarly, standards of coherence could play a role in the strategies that are used. Collectively, the present results indicate that students shift their focus from "learning to read" in Grades 2 and 3 to "reading to learn" in Grades 4 and 5.

We also found that computer reading posed greater distractions than paper reading (e.g., computer characters), which may interfere with control strategies, as others have shown (Panadero, 2017; Salmerón et al., 2021; Storz & Hoffman, 2013). Detrimental effects on reading metacomprehension have been recorded in fifth grade students who use computers (Halamish & Elbaz, 2020), but positive metacognitive effects have been documented in third to fifth grade students who read e-books (Connor et al., 2019). Our findings showed that younger students were more attentive than older students which could be attributed to the nature and content of the reading assignment. In our computer assignment, the stimulating cues and animated characters were simpler in earlier grades but more complex in later grades. This may produce a cognitive load that disables monitoring and control in older students. It is therefore possible that teacher's instruction can aid students in differentiating between those features that promote versus those that impede SRL metacognitive strategies during computer-based reading assignments.

Two findings worth discussing are representation of text in tables and phonic exercises. First, students across grades used embedded charts and tables to transfer content in computer-based reading, but they did not graphically represent content in paper-based reading, with the exception of fifth graders who did. Organizing text graphically aids memorization and comprehension (Ponce et al., 2013; Zimmerman & Martinez Pons, 1986). Our study showed that graphic organizers may contribute to retention but the ability to comprehend was not reported, a finding that contradicts previous research on reading in print (Ackerman & Leiser, 2014; Crooks & Cheon, 2013; Schwartz et al., 1998). A possible explanation for the differences between our findings and those of previous studies is that younger students have not yet learned to recognize patterns in content and meaning, whereas older students have. We suggest that transfer of knowledge (e.g., the ability to graphically organize text) has not been achieved equally well for the two reading formats which may instigate modeling strategies and think alouds from teachers.

Second, students used the phonic approach of sounding words out more in the paper than in the computer condition. However, sounding-out did not yield the desired result (i.e., monitoring comprehension) among students in early grades which is in agreement with previous literature showing that younger students may still be in the process of understanding language conventions (Paris & Flukes, 2005; Ceyhan & Yildiz, 2021). Older students understood which reading comprehension processes were appopriate for paper assignments, but for computer assignments older students relied on electronic characters for phonological processing and phoneme awareness. In fact, the computer medium hindered sounding out, read aloud, and subsequent reading comprehension for students across grades, which is a novel finding in the literature of computer-assisted learning and metacognition. A possible explanation is that students have pobably not received instruction on how to transfer phonic exercises to computer reading assignments. Also, when students had the story narrated to them during the computer condition, this would mean that they were no longer reading which could have influenced the processes involved. Future research should address shortcomings of educational software design and ways teachers could overcome them in eliciting student comprehension and learning.

Finally, students used their performance to evaluate their progress in reading assignments, whether on computer or paper. Our findings in the computer condition showed that our participants received direct and explicit feedback that provided opportunities for students to recheck and correct a response, as past research has shown (Andrade, 2019; Azevedo & Hadwin, 2005; Manlove et al., 2007). This finding may be explained by the fact that computer-based feedback promotes persistence in successfully completing reading assignments. However, multiple rechecking and recalling of information was demonstrated more in the paper- than in the computer-based reading assignments. One possible explanation is that the computer software does not allow students to frequently recheck and revise, whereas the paper worksheet does. These findings suggest that education professionals and software developers should consider equivalent features for

elementary students evaluating, rechecking, and correcting their reading assignments.

Taken together the present and past findings indicate that planning, monitoring, control, and evaluation develop during the elementary school years, and this is slightly more evident in paper- than in computer-based reading assignments.

Implications

Our findings have implications for theory and practice. Our results expand the theoretical perspectives of SRL and metacognition. In this study, elementary grade students exposed to computer-based reading tasks show emerging traits of self-regulation and metacognition as early as second grade, along with increased strategy use and flexibility in the upper elementary grades. The multidimensional aspect of metacognition is comprised of interconnected thought processes and regulatory skills, such as planning, monitoring, and evaluating that motivate learners and improve academic performance (Brown et al., 1981; Flavell, 1979; Pintrich & de Groot, 1990). It appears that elementary students have the potential to apply regulatory metacognitive strategies between reading media.

From a practical standpoint, our results indicate that elementary-aged students can use metacognitive SRL strategies to help themselves complete computer-based reading tasks effectively. Planning, monitoring, control, and evaluation are ways to restore deficiencies in reading while performing computer- and paper-based reading assignments, across grades. However, computer-based reading software can contain confounding features, of which educators need to be aware. For example, the gradual appearance of a passage or feedback from animated characters tends to hinder SRL practices. These elements may interrupt SRL thought process and actions and disable students' own SRL capabilities. Students in computer-based reading assignments can receive instruction to set goals, ask themselves questions, read aloud, perform decoding and phonemic exercises without relying on computer characters that strip independent learning. Students in both conditions can improve cognitive deficiencies by cultivating cognitive schemas (memory, attention) that facilitate connections with different reading genres. Schematic representations using graphic organizers can potentially contribute to forming cognitive connections and applying SRL strategies. Likewise, modeling, and scaffolded instruction can focus on the process of acquiring SRL skills especially in low-stake and non-graded reading assignments. Finally, students across grades can collaborate during computer- and paper-based reading assignments in ways that promote knowledge sharing, peer interaction, and motivate high- and low-skill readers.

Limitations and Strengths

This research had limitations that should be considered. The sample included students from a single school district in a rural area and may not be fully representative of elementary students across the United States. There was unequal gender and grade distribution in the sample due to the school lockdown associated with the COVID-19 crisis. Finally, younger students seemed unfamiliar with the terminology of metacognitive SRL which we overcame by rephrasing question items. This practice, however, may have produced more thorough descriptions of SRL metacognitive processes in younger and older elementary-aged students.

The research also had strengths. Students served as controls of themselves and there was counterbalancing in the presentation of computer- and paper-based reading, which increases confidence that we identified true differences between conditions. Participants were observed and interviewed in a naturalistic setting with typical computer-based and paper-pen reading assignments. We generated in-depth findings by applying triangulation of methods and researchers. Finally, this is one of few studies where direct comparisons in metacognitive SRL processes were made between two popular reading modalities – computer and paper.

Conclusions

According to the qualitative data presented, students in lower and upper elementary grades possess SRL metacognitive skills, which are more readily apparent when performing reading tasks on paper than in computer. The results may imply increased familiarity with reading on paper and use of prior knowledge in the paper condition. Teachers may use these important insights as opportunities to cultivate and transfer SRL metacognitive skills in elementary students between the two reading formats – computer and paper.

References

Ackerman, R., & Leiser, D. (2014). The effect of concrete supplements on metacognitive regulation during learning and open-book test taking. *British Journal of Educational Psychology*, *84*(2), 329–348. https://doi.org/10.1111/bjep.12021

Al-Jarrah, T. M., Mansor, N., Rashid, R. A., Bashir, I., & Al-Jarrah, J. M. (2018). EFL students' attitude toward using metacognitive strategies in writing. *English Language Teaching*, *11*(10), 162–171. https://doi.org/10.5539/elt.v11n10p162

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

Azevedo, R., & Hadwin, A. F. (2005). Scaffolding self-regulated learning and metacognition – Implications for the design of computer-based scaffolds. *Instructional Science*, *33*(5–6), 367–379. https://doi.org/10.1007/s11251-005-1272-9

Baker, S., Elsie, & Edwards, R. (2012). *How many qualitative interviews is enough? Expert voices and early career reflections on sampling and cases in qualitative research* (pp. 1–42). National Centre for Research Methods (NCRM). http://eprints.ncrm.ac.uk/2273/4/how_many_interviews.pdf

Bandura, A. (1991). Social cognitive theory of self-regulation. *Organizational Behavior and Human Decision Processes*, *50*(2), 248–287. https://doi.org/10.1016/0749-5978(91)90022-L

Bannert, M., Sonnenberg, C., Mengelkamp, C., & Pieger, E. (2015). Short-and long-term effects of students' self-directed metacognitive prompts on navigation behavior and learning performance. *Computers in Human Behavior*, *52*, 293–306. https://doi.org/10.1016/j.chb.2015.05.038

Brown, A. L. (1977). *Knowing when, where, and how to remember: A problem of metacognition.* (Technical Report No. 47; Bolt, Beranek and Newman, Inc., Cambridge, Mass., p. 146). Center for the Study of Reading. https://eric.ed.gov/?id=ED146562

Brown, A. L., Campione, J. C., & Day, J. D. (1981). Learning to learn: On training students to learn from texts. *Educational Researcher*, *10*(2), 14–21. https://doi.org/10.3102/0013189X010002014

Bulu, S. T., & Pedersen, S. (2012). Supporting problem-solving performance in a hypermedia learning environment: The role of students' prior knowledge and metacognitive skills. *Computers in Human Behavior*, *28*(4), 1162–1169. https://doi.org/10.1016/j.chb.2012.01.026

Ceyhan, S., & Yildiz, M. (2021). The effect of interactive reading aloud on student reading comprehension, reading motivation and reading fluency. *International Electronic Journal of Elementary Education*, *13*(4), 421–431. https://doi.org/10.26822/iejee.2021.201

Chen, C.-M., Chen, L.-C., & Horng, W.-J. (2019). A collaborative reading annotation system with formative assessment and feedback mechanisms to promote digital reading performance. *Interactive Learning Environments*, 1–18. https://doi.org/10.1080/10494820.2019.1636091

Chen, Q.-S. (2009). Metacomprehension monitoring and regulation in reading comprehension. *Acta Psychologica Sinica*, *41*(8), 676–683. https://doi.org/10.3724/SP.J.1041.2009.00676

Chevalère, J., Cazenave, L., Berthon, M., Martinez, R., Mazenod, V., Borion, M. C., Pailler, D.,

Rocher, N., Cadet, R., Lenne, C., Maïonchi-Pino, N., & Huguet, P. (2021). Computer-assisted instruction versus inquiry-based learning: The importance of working memory capacity. *PLoS ONE*, *16*(11 November). https://doi.org/10.1371/journal.pone.0259664

Clinton, V. (2019). Reading from paper compared to screens: A systematic review and meta-analysis. *Journal of Research in Reading, 42*(2). https://doi.org/10.1111/1467-9817.12269

Combrinck, C., & Mtsatse, N. (2019). Reading on paper or reading digitally? Reflections and implications of ePIRLS 2016 in South Africa. *South African Journal of Education*, *39*. https://eric.ed.gov/?id=EJ1242907

Connor, C. M., Day, S. L., Zargar, E., Wood, T. S., Taylor, K. S., Jones, M. R., & Hwang, J. K. (2019). Building word knowledge, learning strategies, and metacognition with the Word-Knowledge e-Book. *Computers & Education*, *128*, 284–311. https://doi.org/10.1016/j.compedu.2018.09.016

Corbin, J., & Strauss, A. (1990). Grounded Theory Research: Procedures, Canons, and Evaluative Criteria. *Qualitative Sociology*, *13*(1), 1–20.

Corno, L. (1994). Implicit teachings and self-regulated learning. *Invited Address to Divisions K and C*, 52. https://eric.ed.gov/contentdelivery/servlet/ERICServlet?accno=ED377140

Creswell, J. W., & Plano Clark, V. L. (2010). *Designing and conducting mixed methods research* (2nd ed.). SAGE Publications, Inc.

Crooks, S. M., & Cheon, J. (2013). Strategies for note taking on computer-based graphic organizers. In G. Schraw, M. T. McCrudden, & D. Robinson (Eds.), *Learning through visual displays.* (2014-01969-008; pp. 187–221). IAP Information Age Publishing.

Curriculum Associates LLC. (2019). *Personalized learning for all students* [Education & Curriculum]. https://www.curriculumassociates.com/products/i-ready

Dalgarno, B., & Lee, M. J. W. (2010). What are the learning affordances of 3-D virtual environments? *British Journal of Educational Technology*, *41*(1), 10–32. https://doi.org/10.1111/j.1467-8535.2009.01038.x

Deekens, V. M., Greene, J. A., & Lobczowski, N. G. (2018). Monitoring and depth of strategy use in computer-based learning environments for science and history. *British Journal of Educational Psychology*, *88*(1), 63–79. https://doi.org/10.1111/bjep.12174

Delfino, M., Dettori, G., & Persico, D. (2008). Self-regulated learning in virtual communities. *Technology, Pedagogy & Education*, *17*(3), 195–205. https://doi.org/10.1080/14759390802383785

Delgado, P., Vargas, C., Ackerman, R., & Salmerón, L. (2018). Don't throw away your printed books: A meta-analysis on the effects of reading media on reading comprehension. *Educational Research Review* (Vol. 25), 23-38. https://doi.org/10.1016/j.edurev.2018.09.003

Dresel, M., & Haugwitz, M. (2008). A computer-based approach to fostering motivation and self-regulated learning. *Journal of Experimental Education*, *77*(1), 3–18. https://doi.org/10.3200/JEXE.77.1.3-20

Earle, F. S., Del Tufo, S. N., Evans, T. M., Lum, J. A. G., Cutting, L. E., & Ullman, M. T. (2020). Domain-general learning and memory substrates of reading acquisition. *Mind, Brain, and Education*, *14*(2), 176–186. https://doi.org/10.1111/mbe.12234

Ferreira, P. C., Simão, A. M. V., & da Silva, A. I. (2017). How and with what accuracy do children report self-regulated learning in contemporary EFL instructional settings? *European Journal of Psychology of Education*, *32*(4), 589–615. https://doi.org/10.1007/s10212-016-0313-x

Fiorella, L., & Mayer, R. E. (2016). Effects of observing the instructor draw diagrams on learning from multimedia messages. *Journal of Educational Psychology*, *108*(4), 528–546. https://doi.org/10.1037/edu0000065

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

Follmer, D. J., & Sperling, R. A. (2019). Examining the role of self-regulated learning microanalysis in the assessment of learners' regulation. *The Journal of Experimental Education*, *87*(2), 269–287. https://doi.org/10.1080/00220973.2017.1409184

Furenes, M. I., Kucirkova, N., & Bus, A. G. (2021). A comparison of children's reading on paper versus screen: A Meta-Analysis. *Review of Educational Research, 91*(4). https://doi.org/10.3102/0034654321998074

Glaser, B. G., & Strauss, A. L. (1967). *The discovery of grounded theory; strategies for qualitative research*. Chicago: Aldine Pub. Co.

Gray, L., & Lewis, L. (2021). Use of educational technology for instruction in public schools: 2019-20. First look-summary. NCES 2021-017. *National Center for Education Statistics*.

Greene, J. A., Muis, K. R., & Pieschl, S. (2010). The role of epistemic beliefs in students' self-regulated learning with computer-based learning environments: Conceptual and methodological issues. *Educational Psychologist*, *45*(4), 245–257. https://doi.org/10.1080/00461520.2010.515932

Greenhow, C., Graham, C. R., & Koehler, M. J. (2022). Foundations of online learning: Challenges and opportunities. *Educational Psychologist*, *57*(3). https://doi.org/10.1080/00461520.2022.2090364

Groß, D. (2021). In the self-control and self-regulation maze: Integration and importance. *Personality and Individual Differences*, *175*, 110728. https://doi.org/10.1016/j.paid.2021.110728

Guba, E. G., & Lincoln, Y. S. (1994). Competing paradigms in qualitative research. In N. K. Denzin & Y. S. Lincoln (Eds.). In *Handbook of qualitative research* (Vol. 2, p. 105). Thousand Oaks, Sage. http://steinhardtapps.es.its.nyu.edu/create/courses/3311/reading/10-guba_lincoln_94.pdf

Gulek, J. C., & Demirtas, H. (2005). Learning with technology: The impact of laptop use on student achievement. *Journal of Technology, Learning, and Assessment*, *3*(2), 1–39. https://ejournals.bc.edu/index.php/jtla/article/view/1655/1501

Halamish, V., & Elbaz, E. (2020). Children's reading comprehension and metacomprehension on screen versus on paper. *Computers & Education*, *145*, 103737. https://doi.org/10.1016/j.compedu.2019.103737

Hardcastle, J., Herrmann-Abell, C. F., & DeBoer, G. E. (2017). Comparing student performance on paper-and-pencil and computer-based-tests. *ScienceDaily*. https://www.sciencedaily.com/releases/2017/06/170612115723.htm

Howitt, D., & Cramer, D. (2011). *Introduction to research methods in psychology* (3rd ed). Prentice Hall/Pearson.

Kim, H., & Kim, J. (2013). Reading from a LCD monitor versus paper: Teenagers' reading performance. *International Journal of Research Studies in Educational Technology*, *2*(1), 15–24. https://doi.org/10.5861/ijrset.2012.170

Koriat, A. (2012). The relationships between monitoring, regulation and performance. *Learning and Instruction*, *22*(4), 296–298. https://doi.org/10.1016/j.learninstruc.2012.01.002

Kornbluh, M. (2015). Combatting challenges to establishing trustworthiness in qualitative research. *Qualitative Research in Psychology*, *12*(4), 397–414. https://doi.org/10.1080/14780887.2015.1021941

Koutsouraki, S. (2020). Metacognition and reading comprehension: Recent trends in theory, research and practice. *Psychology: The Journal of the Hellenic Psychological Society*, *16*(3), 205–225. https://doi.org/10.12681/psyhps.23815

Krauss, S. E. (2005). Research paradigms and meaning making: A primer. *The Qualitative Report*, *10*(4), 758–770. https://dx.doi.org/10.46743/2160-3715/2005.1831

Kuisma, M., & Nokelainen, P. (2018). Effects of progressive inquiry on cognitive and affective learning outcomes in adolescents' geography education. *Frontline Learning Research*, *6*(2), 1–19. eric. https://doi.org/10.14786/flr.v6i2.309

Latini, N., & Bråten, I. (2022). Strategic text processing across mediums: A verbal protocol study. *Reading Research Quarterly, 57*(2). https://doi.org/10.1002/rrq.418

Lin, C.-H., Zhang, Y., & Zheng, B. (2017). The roles of learning strategies and motivation in online language learning: A structural equation modeling analysis. *Computers & Education*, *113*, 75–85. https://doi.org/10.1016/j.compedu.2017.05.014

Máñez, I., Vidal-Abarca, E., & Magliano, J. P. (2022). Comprehension processes on question-answering activities: A think-aloud study. *Electronic Journal of Research in Educational Psychology*, *20*(56). https://doi.org/10.25115/ejrep.v20i56.3776

Manlove, S., Lazonder, A. W., & Jong, T. (2007). Software scaffolds to promote regulation during scientific inquiry learning. *Metacognition and Learning*, *2*(2–3), 141–155. https://doi.org/10.1007/s11409-007-9012-y

Mayer, R. E. (2003). The promise of multimedia learning: Using the same instructional design methods across different media. *Learning and Instruction*, *13*(2), 125–139. https://doi.org/10.1016/S0959-4752(02)00016-6

McCrudden, M. T., Marchand, G., & Schutz, P. (2019). Mixed methods in educational psychology inquiry. *Contemporary Educational Psychology*, *57*, 1–8. https://doi.org/10.1016/j.cedpsych.2019.01.008

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

Mora, T., Escardíbul, J.-O., & Di Pietro, G. (2018). Computers and students' achievement: An analysis of the One Laptop per Child program in Catalonia. *International Journal of Educational Research*, *92*, 145–157. https://doi.org/10.1016/j.ijer.2018.09.013

Morrow, S. L. (2005). Quality and trustworthiness in qualitative research in counseling psychology. *Journal of Counseling Psychology*, *52*(2), 250–260. https://doi.org/10.1037/0022-0167.52.2.250

Muis, K. R., Psaradellis, C., Chevrier, M., Di Leo, I., & Lajoie, S. P. (2016). Learning by preparing to teach: Fostering self-regulatory processes and achievement during complex mathematics problem solving. *Journal of Educational Psychology*, *108*(4), 474–492. https://doi.org/10.1037/edu0000071

Muis, K. R., Ranellucci, J., Trevors, G., & Duffy, M. C. (2015). The effects of technology-mediated immediate feedback on kindergarten students' attitudes, emotions, engagement and learning outcomes during literacy skills development. *Learning & Instruction*, *38*, 1–13. a9h. https://doi.org/10.1016/j.learninstruc.2015.02.001

Panadero, E. (2017). A review of self-regulated learning: Six models and four directions for research. *Frontiers in Psychology*, *8*(422), 28. https://doi.org/10.3389/fpsyg.2017.00422

Paris, S. G., & Flukes, J. (2005). Assessing children's metacognition about strategic reading. In S. E. Israel, C. C. Block, K. L. Bauserman, & K. Kinnucan-Welsch (Eds.), *Metacognition in literacy learning: Theory, assessment, instruction, and professional development.* (2005-07525-007; pp. 121–139). Lawrence Erlbaum Associates Publishers.

Pintrich, P. R. (1999). The role of motivation in promoting and sustaining self-regulated learning. *International Journal of Educational Research*, *31*(6), 459–470. https://doi.org/10.1016/S0883-0355(99)00015-4

Pintrich, P. R., & de Groot, E. V. (1990). Motivational and self-regulated learning components of classroom academic performance. *Journal of Educational Psychology*, *82*(1), 33–40. https://doi.org/10.1037/0022-0663.82.1.33

Ponce, H. R., Mayer, R. E., & Lopez, M. J. (2013). A computer-based spatial learning strategy approach that improves reading comprehension and writing. *Educational Technology Research and Development*, *61*(5), 819–840. https://doi.org/10.1007/s11423-013-9310-9

Postholm, M. B. (2011). Self-regulated learning in teaching: Students' experiences. *Teachers and Teaching: Theory and Practice*, *17*(3), 365–382.

Pratt, S. M., & Martin, A. M. (2017). The differential impact of video-stimulated recall and concurrent questioning methods on beginning readers' verbalization about self-monitoring during oral reading. *Reading Psychology*, *38*(5), 439–485. https://doi.org/10.1080/02702711.2017.1290726

Pressley, M., & Afflerbach, P. (1995). *Verbal protocols of reading: The nature of constructively responsive reading*. Lawrence Erlbaum Associates.

Price, S., & Oliver, M. (2007). A framework for conceptualising the impact of technology on teaching and learning. *Journal of Educational Technology & Society*, *10*(1), 16–27.

Qi, X. (2021). Effects of self-regulated learning on student's reading literacy: Evidence from Shanghai. *Frontiers in Psychology*, *11*. https://doi.org/10.3389/fpsyg.2020.555849

Robinson, K. (2016). The effect of technology integration on high school student's literacy achievement. *Teaching English with Technology*, *16*(3), 3–16.

Robson, S. (2016). Are there differences between children's display of self-regulation and metacognition when engaged in an activity and when later reflecting on it? The complementary roles of observation and reflective dialogue. *Early Years: An International Journal of Research and Development*, *36*(2), 179–194. https://doi.org/10.1080/09575146.2015.1129315

Roussel, S. (2011). A computer assisted method to track listening strategies in second language learning. *ReCALL*, *23*(2), 98–116. http://dx.doi.org/10.1017/S0958344011000036

Salmerón, L., Delgado, P., Vargas, C., & Gil, L. (2021). Tablets for all? Testing the screen inferiority effect with upper primary school students. *Learning and Individual Differences*, *86*, 8. https://doi.org/10.1016/j.lindif.2021.101975

Saunders, B., Sim, J., Kingstone, T., Baker, S., Waterfield, J., Bartlam, B., Burroughs, H., & Jinks, C. (2018). Saturation in qualitative research: Exploring its conceptualization and operationalization. *Quality & Quantity*, *52*(4), 1893–1907. https://doi.org/10.1007/s11135-017-0574-8

Schacter, D. L., & Szpunar, K. K. (2015). Enhancing attention and memory during video-recorded lectures. *Scholarship of Teaching and Learning in Psychology*, *1*(1), 60–71. https://doi.org/10.1037/stl0000011

Schiavo, G., Mana, N., Mich, O., Zancanaro, M., & Job, R. (2021). Attention-driven read-aloud technology increases reading comprehension in children with reading disabilities. *Journal of Computer Assisted Learning*, *37*(3), 875-886. https://doi.org/10.1111/jcal.12530

Schraw, G., & Dennison, R. S. (1994). Assessing metacognitive awareness. *Contemporary Educational Psychology*, *19*(4), 460–475. https://doi.org/10.1006/ceps.1994.1033

Schunk, D. (2008). Metacognition, self-regulation, and self-regulated learning: Research recommendations. *Educational Psychology Review*, *20*(4), 463–467. https://doi.org/10.1007/s10648-008-9086-3

Schwartz, N. H., Ellsworth, L. S., Graham, L., & Knight, B. (1998). Accessing prior knowledge to remember text: A comparison of advance organizers and maps. *Contemporary Educational Psychology*, *23*(1), 65–89. https://doi.org/10.1006/ceps.1997.0958

Serrano, M.-Á., Vidal-Abarca, E., & Ferrer, A. (2018). Teaching self-regulation strategies via an intelligent tutoring system (TuinLECweb): Effects for low-skilled comprehenders. *Journal of Computer Assisted Learning*, *34*(5), 515–525. https://doi.org/10.1111/jcal.12256

Sha, L., Looi, C.-K., Chen, W., Seow, P., & Wong, L.-H. (2012). Recognizing and measuring self-regulated learning in a mobile learning environment. *Computers in Human Behavior*, *28*(2), 718–728. https://doi.org/10.1016/j.chb.2011.11.019

Smith, M. (2016). Computer science for all. *Whitehouse.Gov*. https://obamawhitehouse.archives.gov/blog/2016/01/30/computer-science-all

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*(1), 51–79. https://doi.org/10.1006/ceps.2001.1091

Squire, L. R. (2009). Memory and brain systems: 1969-2009. *Journal of Neuroscience, 29*(41), 12711-12716. https://doi.org/10.1523/JNEUROSCI.3575-09.2009

Stoop, J., Kreutzer, P., & Kircz, J. (2013). Reading and learning from screens versus print: A study in changing habits: Part 1 – reading long information rich texts. *New Library World*, *114*(7/8), 284–300. https://doi.org/10.1108/NLW-01-2013-0012

Storz, M. G., & Hoffman, A. R. (2013). Examining response to a one-to-one computer initiative: Student and teacher voices. *RMLE Online: Research in Middle Level Education*, *36*(6), 1–18. https://doi.org/10.1080/19404476.2013.11462099

Tashakkori, A., & Teddlie, C. (1998). *Mixed methodology: Combining qualitative and quantitative approaches* (Vol. 46). SAGE Publications.

The White House. (2021). *Executive order on supporting the reopening and continuing operation of schools and early childhood education providers* [Governmental]. The White House Briefing Room. https://www.whitehouse.gov/briefing-room/presidential-actions/2021/01/21/executive-order-supporting-the-reopening-and-continuing-operation-of-schools-and-early-childhood-education-providers/

U.S. Department of Education. (2016). *Future ready learning: Reimagining the role of technology in education*. Office of Educational Technology. https://tech.ed.gov/netp/learning/

Vidal-Abarca, E., Mañá, A., & Gil, L. (2010). Individual differences for self-regulating task-oriented reading activities. *Journal of Educational Psychology*, *102*(4), 817–826. https://doi.org/10.1037/a0020062

Wilson, V. A. (1998). *Qualitative research: An introduction. Purposes, methodology, criteria for judgment, and a rationale for mixed methodology.* (Mixed Method Evaluation ED 423 285; p. 18). https://files.eric.ed.gov/fulltext/ED423285.pdf

Xu, Y., Wang, D., Collins, P., Lee, H., & Warschauer, M. (2021). Same benefits, different communication patterns: Comparing children's reading with a conversational agent vs. a human partner. *Computers & Education*, *161*, 104059. https://doi.org/10.1016/j.compedu.2020.104059

Yin, R. K. (2011). Ch. 5 Doing Fieldwork. In *Qualitative research from start to finish* (pp. 109–128). Guilford Press.

Zheng, B., Warschauer, M., Lin, C.-H., & Chang, C. (2016). Learning in one-to-one laptop environments: A meta-analysis and research synthesis. *Review of Educational Research*, *86*(4), 1052–1084. https://doi.org/10.3102/0034654316628645

Zimmerman, B. J. (1989). A social cognitive view of self-regulated academic learning. *Journal of Educational Psychology*, *81*(3), 329–339. https://doi.org/10.1037/0022-0663.81.3.329

Zimmerman, B. J., & Martinez Pons, M. (1986). Development of a structured interview for assessing student use of self-regulated learning strategies. *American Educational Research Journal*, *23*(4), 614–628.

iejee

International Electronic Journal of Elementary Education
September 2023, Volume 16, Issue 1, 17-30

Examining Students' Proof Writing and Justification Skills in the Context of Sum of Measures of Polygons' Interior Angles

Handan Demircioğlu[a,*], Kudret Hatip[b]

Received : 17 December 2022
Revised : 15 July 2023
Accepted : 29 September 2023
DOI : 10.26822/iejee.2023.311

[a,*] **Corresponding Author:** Handan Demircioğlu, Faculty of Education, Sivas Cumhuriyet University, Sivas, Turkey.
E-mail: handandemircioglu@gmail.com
ORCID: https://orcid.org/0000-0001-7037-6140

[b] Kudret Hatip, Ministry of Education, Sivas Turkey
E-mail: hatipkudret@gmail.com
ORCID ID: https://orcid.org/0000-0002-5103-4759,

Abstract

The present study aims to examine 8th grade students' proof writing and justification skills. The research was conducted using the document analysis method. The participants of the study consisted of 16 voluntaries 8th grade students. The participants were determined according to the convenience sampling method. Data were collected with the "geometric proof writing and justification test" prepared by the researchers. The data collection tool was prepared by referring to the 5th, 6th, and 7th grade curricula. "Proof Writing Rubric" and "Justification Rubric" were used to analyze the data. The results of the study showed that 8th grade students' geometric proof writing and justification skills were at low levels. In addition, it was revealed that their justification skill levels were lower than their proof writing levels.

Keywords:

Proof, Justification Skills, Mathematics Education

Introduction

The mathematical thinking process includes high-level thinking skills such as specialization, generalization, prediction, generating assumptions, and checking the accuracy of assumptions (Mason, Burton & Stacey, 2010). These skills are closely related to proof skills. Proof skill includes understanding the proof of a mathematical statement in addition to being able to recognize and justify the construction of the proof. The importance of mathematical thinking is frequently emphasized in all learning areas of mathematics. Geometry is one of these learning areas. Difficulties and deficiencies in geometry teaching and students' lack of success in geometry learning are among the most emphasized problems by educators (Alex & Mammen, 2012). Geometry teaching is supposed to contribute to the development of students' ability to visualize objects in their minds, reducing the objects they encounter in daily life to two dimensions, solving problems, making assumptions, making logical inferences, and making proofs. Generalization, reasoning, and justification (Ministry of National Education [MoNE], 2020) and the process of constructing geometric ideas in a meaningful way are among the skills that students should acquire in geometry

teaching (Driscoll, DiMatteo, Nikula & Egan, 2007). Despite this importance, studies show that students not only have difficulty in understanding rules and operations in the geometry learning domain but also have difficulty explaining their solutions and ways of thinking, understanding proof and writing proof (Almeida, 2000; Jones, 2000; Hadas, Hershkovitz & Schwarz, 2000, Stylianides, Stylianides, & Philippou, 2004; 2007).

Although proof and justification skills have an important place in mathematics in general and geometry in particular, there are limited studies on proof writing and justification (Coşkun, 2019; Dimakos, Nikoloudakis, Ferentinos & Choustulakis, 2007; Senk, 1983;1985;1989; Şen & Güler, 2022, Özmusul, 2018). Senk (1985) stated that, considering the current curriculum and typical teaching practices, students do not master the skills required in a standard geometry course, and even students could not realize the necessity of proof while proving in geometry and were insufficient in proof writing types. Senk (1989) examined the relationships between van Hiele levels, success in writing geometry proofs, and success in geometry. The results of the study showed that students' success in geometry proof writing was positively related to van Hiele's level of geometric thinking and success in standard geometry course. Dimakos, Nikoloudakis, Ferentinos and Choustulakis (2007) revealed that students have difficulties in proof writing in geometry and even they do not know how to start making proofs. Özmusul (2018) examined the levels of justification. The results of the study showed that the justification skills of the participants were low and the students with high achievement levels had high justification skills. Harel and Sowder (1998) made a classification of proof based on the arguments and justifications used by students to explain the decision (correctness or falsity of the statement) about a mathematical statement. In the light of the studies, it is seen that there is a great relationship between proof and justification. Considering these studies, the results of a study in which proof writing and justification skills in geometry are examined together will contribute to the literature. Therefore, the aim of this study is to examine the 8th grade students' proof writing and justification skills in the context of the sum of the measures of the interior angles of polygons and to compare the answers of the students with the lowest and highest skills in both proof writing and justification. In this way, geometric proof writing and justification skills will be evaluated together and contribute to the literature.

Conceptual Framework

Proof writing

Proof has many functions in mathematics education. The most basic role of proof is to show that a claim is correct or incorrect. Students generally perceive the concept of proof in this sense. For many students, the proof is a practice that needs to be memorized by the teacher with standard methods and steps. Knuth (2002) stated that teachers tend to view proof in a pedagogically limited way, that is, as a subject of study rather than a tool for communicating and studying mathematics. According to mathematics educators, proof is a method of thinking as well as an important skill used to explain why a claim is correct. The proof is categorized into two main categories. The first is proofs that show the correctness of a claim superficially and do not involve too much questioning. The second one is explanatory proof that responds to questions of why and why and reveals the correctness of the claim in depth (Bayazıt, 2017). Almeida (2000) emphasizes that understanding proof and writing proof is one of the main distinguishing features of mathematics. Goetting (1995) stated that there are three different understandings of proof in his study in which they examined their understanding of proof - the arguments they found persuasive and the arguments they accepted as valid evidence. These insights are not necessarily imprecise, in the form of a supporting argument, a means to conclusively confirm assumptions, a statement where precise verification is necessary but sometimes not sufficient, or a classroom exercise.

The results of studies on proof writing in geometry showed that students' proof writing skills were weak and they had difficulties (Daguplo, 2014; Güner & Topan, 2016; Şen & Güler, 2022). Güner and Topan (2016) revealed that primary school mathematics teacher candidates have weak geometric proof skills, they have difficulty in proving, they have misconceptions that a single example or numerical representation showing accuracy is sufficient for proof, and they cannot transfer their existing knowledge to the proof process. Şen and Güler (2022) examined the effectiveness of teaching activities based on the Van Hiele model on geometric proof writing skills. The results of the study showed that teaching activities based on the Van Hiele model supported the development of pre-service teachers' proof writing skills. Daguplo (2014) measured students' geometry proof writing performance according to Van Hiele's geometric thinking model levels. At the end of the study, it was seen that the pre-service teachers were not at the highest level in terms of proof writing. In summary, it is seen that they have difficulties in proofs writing in geometry. One of skills that is effective in the development of proof writing skill is justification.

Justification

Justification skill is as important as proof skill in mathematics teaching. There are different definitions of justification in the literature. National Council of Teachers of Mathematics (NCTM) (1989) stated

that mathematics is a set of justifications. Likewise, justification is emphasized in the special objectives set by MoNE (2020). In line with these specific objectives, students will be able to express their thoughts and reasoning in the problem-solving process and see the gaps or deficiencies in the mathematical reasoning of their classmates. The realization of these objectives depends on the development of students' justification skills. As Ross (1998) states, if students' reasoning skills are not developed, mathematics will remain a mass calculation without thinking.

Different forms of justification have been defined in the literature by examining students' justification skills (Balacheff, 1988; Bell, 1976; Harel & Sowder, 1998; Marrades & Gutiérrez, 2000). For example, Bell (1976) divided mathematical justification into two categories. The first is "empirical" justification in which the correctness of a mathematical statement is demonstrated with the help of examples. The second is "deductive" justification in which inferences are used in connection with the results. Marrades and Gutiérrez (2000) classified these two categories in more detail. Balacheff (1988) divided the types of justification into two groups. The first group is the "pragmatic justification", which is based on the use of examples or demonstrations. The second is the "conceptual justification", which is based on conceptual facts, abstract formulas, and relationships between properties of mathematical expressions. Harel and Sowder (1998) grouped the justifications used by students to explain the correctness (or incorrectness) of a mathematical statement into three categories as "External Sources-Based, Empirical, and Analytical". On the other hand, justification is discussed in various aspects when the studies in the literature are examined. Justification in the process of generalizing patterns (Akkan, Öztürk & Akkan, 2017; Tanışlı, Yavuzsoy Köse & Camci, 2017; Lannin, 2005); justification in the problem-solving process (Akkuş, 2019), the relationship between justification skills and other variables (Özmusul, 2018) are among these studies. Özmusul (2018) examined whether 7th grade students' justification skills differed according to gender, school, and achievement test score. It was stated that the participants' complete and persuasive justification skills were low, while the justification skills of students with high achievement levels were high. In this study, justification was accepted as correct reasoning to support arguments. According to the arguments put forward in proof writing, it has been discussed as "complete justification" and "partial justification".

The sum of the measures of the polygons' interior angles

The present study's context is "the sum of the measures of the polygons' interior angles". Naturally, students are expected to generalize the reasoning they do in special cases such as triangles and quadrilaterals to polygons while proving and justifying a polygon. Activities related to the topic "Angles in Polygons" are included in the 7th grade textbooks. For example, in the activity "calculating the sum of the interior angles of a polygon" in the MoNE textbook (MoNE, 2014; p.125), while finding the sum of the measures of a polygon's interior angles, the given polygon was divided into triangular parts. The sum of the measures of the polygon's interior angles was found by using the sum of the measures of the polygon's interior angles. In the next activity, the student was asked to find the sum of the measures of the pentagon's interior angles by drawing and explaining the reasons. This way, they are expected to experiment with different polygons. It tried to make students realize that polygons can be divided into triangular regions with diagonals in this activity. Then, a table was created in which the number of these triangular regions was related to the number of sides of the polygons. With the help of this table, students were expected to calculate the sum of the interior angles of polygons with the help of the triangles formed. As can be seen in this activity, the basic skill expected to be developed in students in the context of the sum of measures of the interior of polygons in particular, and geometry teaching, in general, is the ability to reason and justify (NCTM, 1989; MoNE, 2018). In this context, the problem of this study is "What is the relationship between justification and proof writing skills?"

Method

In this study, document analysis was used. Document analysis is the analysis of written materials containing information about the phenomena aimed to be researched (Yıldırım & Şimşek, 2016). Document analysis is used as a solo research method, especially in cases where direct interviews and observations are not possible. This method involves the analysis of written and oral materials containing information about the subjects planned to be researched. Document analysis includes the analysis of written materials containing information about events or phenomena that are intended to be researched. The document review conducted in our research covers the analysis of written materials containing information about the subjects planned to be researched.

Participants

The participants were determined as 8th grade students since it was thought that it would allow the best explanation of the researched topic and provide the best contribution to the solution of the research problem. The participants of the study were 16 volunteer students who continued their education in the 8th grade in the 2020-2021 academic year in a private school in the center of Sivas and participated

in the face-to-face teaching process on the day of data collection. Participants were determined according to the convenience sampling method. Seven of these students were female and nine were male. Participants were named S1, S2, S3....

Data Collection Tool

The data collection tool developed to examine 8th grade students' geometric proof writing and justification skills consist of three questions. Data collection tool questions are from specific to general. The students were asked to prove and justify the sum of the measures of the interior angles of a triangle, quadrilateral, and any n-sided polygon. To enable students to do the proof and justification separately, each question consists of two parts option a and option b.

Table 1.
Data collection tool questions

Sum of interior angles measures	1) a) Demonstrate that the sum of the interior angles of a triangle is 180°. b) What can you say to convince others that your result is correct? Explain.
	2) a) Demonstrate that the sum of the interior angles of a quadrilateral is 360°. b) What can you say to convince others that your conclusion is correct? Explain.
	3) a) Demonstrate that the formula for the sum of the interior angles of a polygon with "n" sides is (n-2).180°. b) What can you say to convince others that your result is correct? Explain.

Data Collection

The data were collected by the researchers in a class hour from the 8th grade students who attended a face-to-face mathematics lesson in a private school in Sivas (due to the pandemic conditions). As it was intended to reflect the current situation, the students were not given correction or justification training before the data of the study were collected. The researchers only gave information about how the test should be done during the data collection process. The students' answers were collected in writing. There was no time limit.

Data Analysis

In analyzing the questions in the data collection tool, two rubrics developed by the researchers were used. The first one is the proof writing rubric and the second one is the justification rubric. While creating these rubrics, studies in the related literature (Coşkun, 2009; Senk, 1983; Sowder & Harel, 1998; Özmusul, 2018) were examined. The proof writing rubric is given in Table 2 and the justification rubric is given in Table 4. After the preparation of the rubrics, the opinions of three mathematics educators were obtained. Miles and Huberman's (1994) reliability formula [Reliability=Agreement/ (Agreement + Disagreement)] was used and the compliance rate was determined as 90%.

Table 2.
Proof Writing Rubric

Criteria	Score
Situations where the solution was completely incorrect, the problem is not understood, or nothing is done.	0
Situations where the question is understood. In other words, the question is expressed verbally, the algebraic form of the question is written or short notes are taken about this expression, a graph is drawn, a table is created, and the expression/truth of the given argument/ proposition is explained with examples.	1
Situations where the question is comprehended. That is, they understand exactly what needs to be proved, determined the method of proof, and created/realized the logical steps given for this, but can not fully conclude the proof or there are deficiencies/errors in some stages of the proof.	2
Situations where the proof is properly completed in the correct form.	3

As can be seen in Table 2, there is one point for each situation in the proof writing section. If the student does not answer any of the three questions in this test, the student gets a total of zero points and if the student answers all of them correctly, the student gets a total of nine points. Table 3 shows the students' level of proof writing according to their scores.

Table 3.
Proof Writing level

Proof writing level	Proof writing score
Unsuccessful	0, 1,2,3
Moderately successful	4,5,6
Very successful	7,8,9

As can be seen in Table 3, if a student's total proof writing score is three or less than three, the student is considered to be "unsuccessful" in proof writing, if the student scores four, five and six points, the student is considered to be "moderately successful" and if the student scores more than six points, the student is considered to be "very successful". The scoring of the justification rubric consists of four supercodes and seven codes.

Table 4.
Justification Rubric

Codes	Supercodes	Score
Answers proving the question correctly and supporting it with correct mathematical reasoning	Complete justification	3
Proving the question correctly but writing the reasoning incompletely	Partial justification	2
Incomplete proof of the question and incomplete writing of the justification		
Writing the appropriate justification for the answer in the questions where the proof is proved incorrectly, a calculation error is made or concept errors are made	Incorrect justification	1
The proof is correct but no justification	Not writing justification	0
The proof is incorrect and no justification		
The proof is missing and no justification		

As can be seen in Table 4, when evaluating the students' justification skills, if they write a correct justification in a question, they receive three points, and if they fail to write a correct justification, they receive zero points. That is, if the student does not answer any of the three questions in the data collection tool, the student receives a total of zero points, and if the student answers all of them correctly, the student receives a total of nine points. The evaluation criteria in the rubric are the same as the rubric for writing proofs. The justification levels of the students according to their scores are given in Table 5.

Table 5.
Level of justification

Level of justification	Justification score
Unsuccessful	0, 1,2,3
Moderately successful	4,5,6
Very successful	7,8,9

As can be seen in Table 5, if a student's total justification score is three or less than three, the student is considered to be "unsuccessful" in justification, if the student scores four, five and six points, the student is considered to be "moderately successful", and if the student scores more than six points, the student is considered to be "very successful".

Results

The student's ability to write proofs and justification skills related to the sum of the measures of the polygons' interior angles were first given in question-by-question tables with the scores they received. Afterward, for each question, excerpts from the answers written by the students were given. In this way, both the levels of proof writing, justification, and generalization were examined in detail. Table 6 shows the students' proof writing and justification skills scores.

Table 6.
Proof Writing and Justification Scores

Student	Triangle		Quadrilateral		Polygon		Total	
	Justification	Writing proof	Justification	Writing proof	Justification	Writing proof	Justification	Writing proof
S1	2	3	0	2	3	3	5	8
S2	3	3	3	3	3	3	9	9
S3	0	1	0	1	0	0	0	2
S4	0	1	0	1	3	3	3	5
S5	0	2	0	1	0	0	0	3
S6	0	0	0	0	0	0	0	0
S7	3	2	2	2	0	0	5	4
S8	0	1	0	1	0	0	0	2
S9	2	2	0	0	2	2	4	4
S10	0	1	0	1	0	0	0	2
S11	0	0	0	1	0	0	0	1
S12	0	1	0	1	1	2	1	4
S13	2	1	0	0	2	2	4	3
S14	0	1	2	2	0	0	2	3
S15	0	3	2	1	1	3	3	7
S16	0	1	0	1	0	0	0	2

As seen in Table 6, 11 students' justification skills scores are three or below three. Therefore, these students were coded as unsuccessful. One student (S2) received a full score and four students (S1, S7, S9, and S13) received four and five scores, so it was seen that these students were "moderately successful". Nine students had a proof writing score of three or less three and were found to be unsuccessful. One student (S2) received nine full scores. Besides, three students (S1, S2, and S15) scored seven or more points and were "very successful". Four students (S4, S7, S9, and S12) scored between three and six points and were "moderately successful". The findings obtained from both skills are summarized in Table 7.

As seen in Table 7, the justification levels of students who had a very good level of proof (S1, S2, and S15) were very good, moderate, and unsuccessful. The justification levels of the students (S4, S7, S9, and S12) who had a moderate level of proof were moderate or unsuccessful. It can be said from these findings that the proof writing levels of the students participating in the study were higher than their justification levels. Only the S13 justification skill level was higher than the level of writing proof. In the next section, excerpts from the student's answers to each question are given. On the other hand, when proof writing and justification skills are examined in the generalization process, that is, in the context of the sum of the interior of the triangle, quadrilateral, and polygon, the highest total score in proof writing was obtained in the sum of the interior angles of a triangle, as can be seen from Table 8. In the sum of the measures of the interior angles of quadrilateral and polygon, the sum of the scores is equal. In justification skills, the highest total score was obtained in the sum of the internal angles of a polygon. After that, in terms of total points, it was obtained in the sum of the measures of the interior angles of the triangle and then the quadrilateral. In this sense, when analyzed in terms of the total scores obtained, no parallelism was found in terms of proof writing and justification skills. That is, a high score in proof writing skills did not require a high score in justification skills.

Considering the scores obtained in the generalization process from triangle to polygon, it is difficult to say that there is an order.

Table 7.

Participants' levels of writing and justifying proofs

Student	Proof writing level	Justification level
S1	Very good	Moderate
S2	Very good	Very good
S3	Unsuccessful	Unsuccessful
S4	Moderate	Unsuccessful
S5	Unsuccessful	Unsuccessful
S6	Unsuccessful	Unsuccessful
S7	Moderate	Moderate
S8	Unsuccessful	Unsuccessful
S9	Moderate	Moderate
S10	Unsuccessful	Unsuccessful
S11	Unsuccessful	Unsuccessful
S12	Moderate	Unsuccessful
S13	Unsuccessful	Moderate
S14	Unsuccessful	Unsuccessful
S15	Very good	Unsuccessful
S16	Unsuccessful	Unsuccessful

Table 8.

Proof writing and justification skills in the generalization process

	Justification				Proof writing			
Student	Triangle	Quadrilateral	Polygon	Total	Triangle	Quadrilateral	Polygon	Total
S1	2	0	3	5	3	2	3	8
S2	3	3	3	9	3	3	3	9
S3	0	0	0	0	1	1	0	2
S4	0	0	3	3	1	1	3	5
S5	0	0	0	0	2	1	0	3
S6	0	0	0	0	0	0	0	0
S7	3	2	0	5	2	2	0	4
S8	0	0	0	0	1	1	0	2
S9	2	0	2	4	2	0	2	4
S10	0	0	0	0	1	1	0	2
S11	0	0	0	0	0	1	0	1
S12	0	0	1	1	1	1	2	4
S13	2	0	2	4	1	0	2	3
S14	0	2	0	2	1	2	0	3
S15	0	2	1	3	3	1	3	7
S16	0	0	0	0	1	1	0	2
Total	12	9	15		23	18	18	

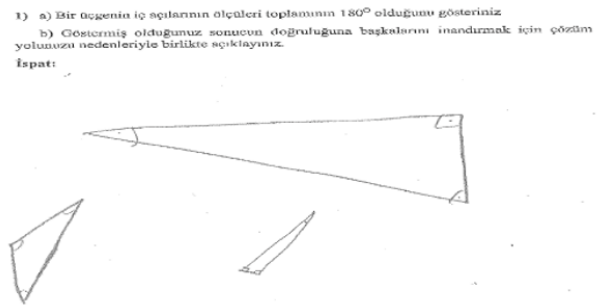
Findings Obtained from the sum of the measures of the triangle's interior angles

Students were asked to prove that the sum of the measures of the triangle's interior angles is 180o and to write their reasons. Here are examples of the answers of the students with the lowest scores and the answers of the students with the highest scores. S6 and S11 got zero points in proof writing and S3, S4, S5, S6, S8, S10, S11, S12, S14, S15, S16 got zero points in justification. The proof of S6, who got zero points in both proof writing and justification, is given in Figure 1a and the justifications are given in Figure 1b.

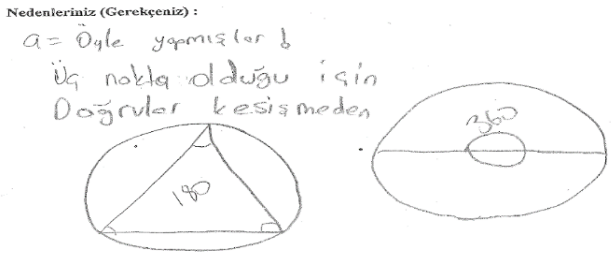
Figure 1.

S6's proof and justifications for the sum of the triangle's interior angles

a)



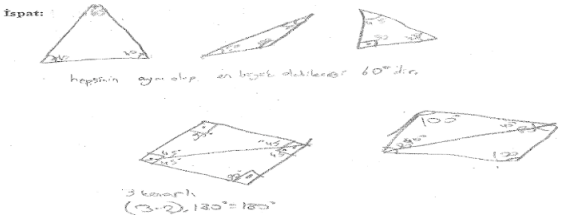
b) Transcription: they did so.. because there are three points...without crossing lines



When Figure 1a is analyzed, S6 drew three different triangles but did not write what the student wanted to explain. According to the proof evaluation rubric, this student's answer was given zero points. S6's justification was neither compatible with the proof nor did the student make a connection with the fact that the sum of the measures of the triangle's interior angles was 180o. Therefore, zero points were given. The students who got one point in this question in writing proof were S3, S4, S8, S10, S12, S13, S14, and S16. According to the justification skills rubric, no student scored one point. That is, no answer fulfills the criterion of not writing a justification appropriate to the answer in the questions where the proof is proved incorrectly, an operation error is made or concept errors are made. The answer of S12 who got one point in this question is given in Figure 2.

Figure 2.

S12's proof for the sum of the measures of the triangle's interior anglesv



Transcription: They are all the same and the maximum they can be is 60 degrees.

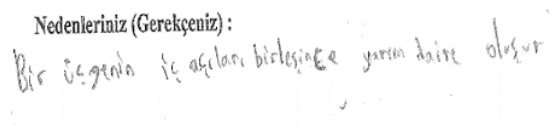
When Figure 2 is analyzed, S12 followed two different ways. In one of them, the student drew different triangles and completed 180o by giving different values to the measures of their interior angles. In the other one, the student drew different quadrilaterals and divided them into two and actually understood the question but could not prove it completely. One point was given according to the proof rubric. S5, S7, and S9 were the students who got two points in proof writing. The proof of S7 is given in Figure 3a. The students who got two points according to the justification rubric were S9 and S13. S9's justifications are given in Figure 3b.

Figure 3.

S7's proof for the sum of the measures of the triangle's interior angles and S9's justifications

a)

b) Transcription: When the interior angles of a triangle converge, a semicircle is formed.

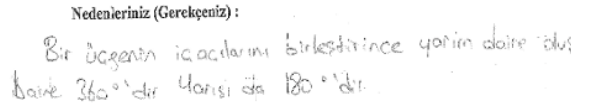


As seen in Figure 3a, S7 drew a triangle and made a right angle by connecting the angles but did not name the angles. Therefore, two points were given because it was not understood which angle was connected where and how in the right angle. As seen in Figure 3b, S9's justification is not clear. The student did not explain what the semicircle formed when the interior angles of the triangle meet. Therefore, two points were given. S1, S2, and S15 in writing proofs, and S2 and

S7 received three points according to the justification skills rubric. The proof of S1 is given in Figure 4a and the justification of S7 is given in Figure 4b.

Figure 4.
S1's proof for the sum of the measures of the triangle's interior angles and S7's justifications
a)

b) Transcription: When the interior angles of a triangle are combined, a semicircle is formed; the circle is 360o, half of which is 180o

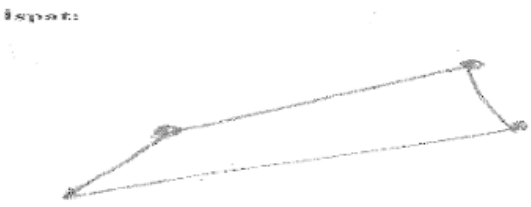


As can be seen in Figure 4a, S1 drew the triangle and named the angles, made a right angle by combining them, and got three points for proving the sum of the measures of the triangle's interior angles completely. S7, on the other hand, stated that a semicircle was formed when the interior angles of the triangle were joined and stated that the circle was 360o and half of it was 180o and wrote the justification by the proof.

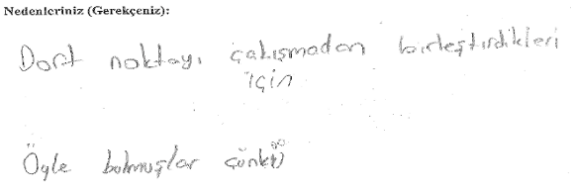
Results obtained from the sum of the measures of the quadrilateral's interior angles

The students were asked to prove that the sum of the measures of the interior angles of a quadrilateral is 360o and to write their reasons. Here are examples of the answers of the students with the lowest scores and the answers of the students with the highest scores. In writing proofs, S6, S9, and S13 received zero points. In justification skills, S1, S3, S4, S5, S6, S8, S9, S10, S11, S12, S13, S16 received zero points. The proof writing answer of S6 is given in Figure 5a and the justifications of S8 are given in Figure 5b.

Figure 5.
S6's proof for the sum of the measures of the interior angles of a quadrilateral and S8's justifications
a)

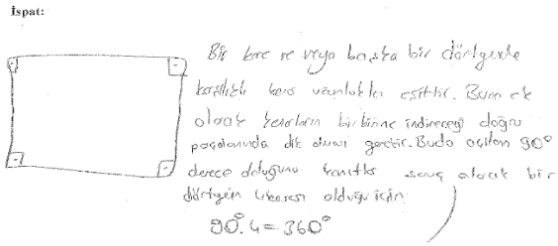


b) Transcription: Since they connect the four points without overlapping. Because that's just as they found it.



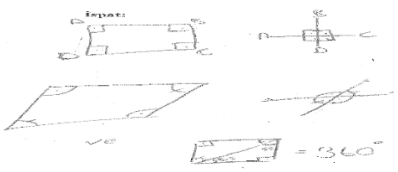
As can be seen in Figure 5a, S6 drew only a quadrilateral and was given zero points because the student did not show what s/he was trying to prove or what s/he wanted to explain on the quadrilateral. S8 (Figure 5b) could not present anything as justification. In writing proofs, S3, S4, S5, S8, S10, S11, S12, S15 and S16 received one point. The answer to S3 for writing proof is given in Figure 6 No student got one point in justification skills. That is, in the questions in which the proof was proved incorrectly, operation errors made, or concept errors were made, there was no answer by the criterion of not writing an appropriate justification for the answer.

Figure 6.
S3's proof for the sum of the measures quadrilateral's interior angles

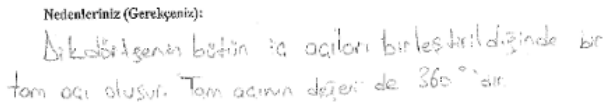


As can be seen in Figure 6, S3 understood the question but did not write an answer to the criteria for writing proofs. The student showed only squares and rectangles and said that one interior angle was 90o and there were four of them, thus s/he showed that the sum of the measures of the interior angles was 360o. However, since what the student wrote was an acceptance and his/her proof would not be accepted for quadrilaterals other than square and rectangle such as rhombus, parallelogram, etc., the student was given one point. S1, S2, S7, and S14 received two points in proof writing and S7, S14, and S15 received two points in justification skills. S1's answer for writing proof is given in Figure 7a. S7's justification answer is given in Figure 7b.

Figure 7.
S1's proof for the sum of the measures of the interior angles of a quadrilateral and S7's justifications
a)

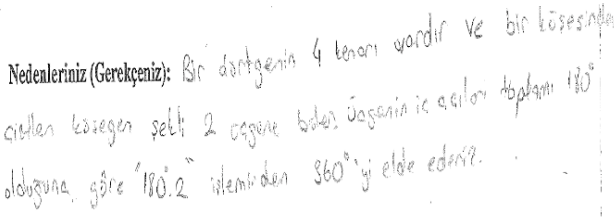


b) Transcription: When all the interior angles of a quadrilateral are combined a complete angle is formed the measure of a full angle is 360o



As can be seen in Figure 7a, S1 drew different quadrilaterals and joined their corners. Then, the student drew a diagonal in a quadrilateral and divided the shape into two. If the student had tried to prove one of the things s/he did and made explanations in that direction, s/he could have got full marks. S7, on the other hand, made his/her justifications only on rectangles, so s/he was given two points. No student got three points in writing proofs. The answer of S2 who got three points in justification skills is given in Figure 8.

Figure 8.
S2's justifications



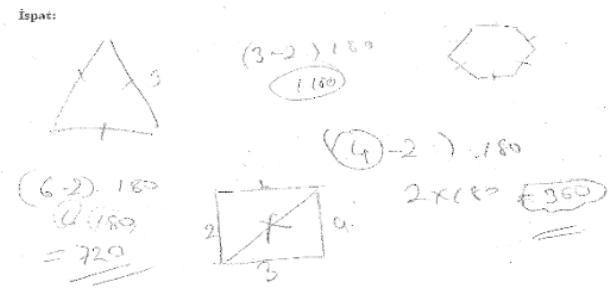
Transcription: A quadrilateral has four sides and the diagonal drawn from one corner divides the shape into two triangles, since the sum of the interior angles of the triangle is 180o, we get 360o from 2.180

As can be seen in Figure 8, S2 made a generalization for all quadrilaterals and wrote that the diagonal drawn from a corner of a quadrilateral divides the quadrilateral into triangles and the sum of the measures of a triangle's interior angles is 180o and the sum of the measures of the interior angles of quadrilaterals is 360o when there are two of them. Therefore, the student received three points.

Results Obtained from the sum of the measures of any polygon's interior angles

Students were asked to show that the sum of the measures polygon's interior angles with n sides is (n-2). 180o and to write their reasons. Here are examples of the answers of the students with the lowest scores and the answers of the students with the highest scores. In writing proofs, S3, S5, S6, S7, S8, S10, S11, S14 and S16 received zero points. In justification skills, S3, S5, S6, S7, S8, S10, S11, S14 and S16 received zero points. S10's proof writing answer is given in Figure 9a and his/her justification is given in Figure 9b.

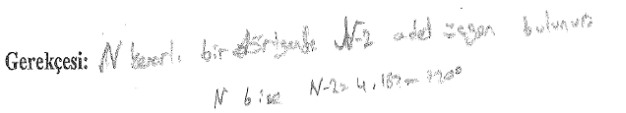
Figure 9.
S10's proof and justifications for the sum of the measures of any polygon's interior angles
a)



b) Transcription: Trial and error

As can be seen in Figure 9a, S10 drew the shapes, and found the sum of the measures of the interior angles, but did not prove what was asked. Therefore, zero point was given. S10 only wrote "trial and error" in the justification. This was not accepted as a justification. No student received a point for writing proof. That is, no answer met the conditions that the question was understood (the question was expressed verbally, the algebraic form of the question was written or short notes were taken about this expression, a graph was drawn, a table was created, the expression/ correctness of the given argument/proposition was tested with examples). In justification skills, S12 and S15 got one point. S12's justification is given in Figure 10.

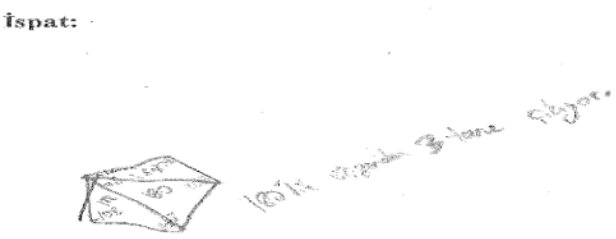
Figure 10.
S12's justification answer



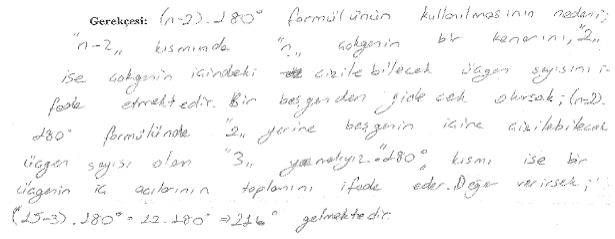
Transcription: A quadrilateral with n sides has n-2 triangles. If n=6, then n-2=4 and 4.180=720.

S12 showed the number of triangles formed by the diagonals drawn from a corner of a pentagon in Figure 11. Here, the student wrote a quadrilateral with n sides. Moreover, he got one point for writing that (n-2) is the number of triangles. S9, S12, and S13 received two points in proof writing and S9 and S13 received two points in justification skills. S12's proof writing answer is given in Figure 11a and S13's justifications are given in Figure 11b.

Figure 11.
S12's proof writing and S13's justifications
a)



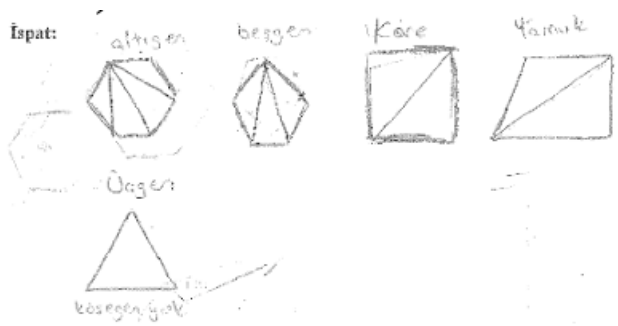
b)



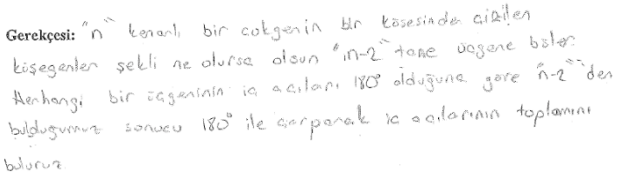
Transcription: The reason for using the formula (n-2)180 is that in the n-2 part, n represents one side of the polygon, and 2 represents the number of triangles that can be drawn inside the polygon. In the (n-2)180 formula, instead of 2, we should write 3, the number of triangles that can be drawn inside the pentagon. The 180 part represents the sum of the interior angles of a triangle. If we give the value (15-3).180=12.180=216.

As can be seen in Figure 11a, S12 drew only one pentagon and started from the number of triangles formed in this pentagon. However, two points were given for drawing only a pentagon. S13 stated the number n as the number of sides of a polygon. But instead of (n-2) being the number of triangles, he said that two is the number of triangles. Therefore, two points were given. In writing proofs, S1, S2, S4, and S15 received three points. The students who got three points in justification skills were S1, S2, and S4. S2's proof is given in Figure 12a and his justifications are given in Figure 12b.

Figure 12.
S2's proof and justifications for the sum of the measures of the interior angles of any polygon
a)



b)



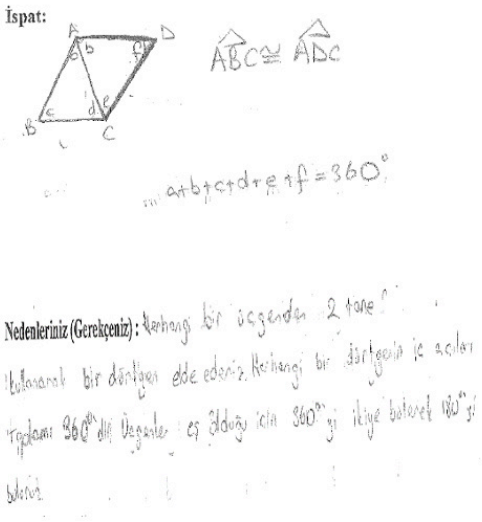
Transcription: The diagonals drawn from one corner of an n-sided polygon divide it into n-2 triangles, regardless of their shape. Since the interior angles of any triangle are 180o, we find the sum of the interior angles by multiplying n-2 by 180o.

As seen in Figure 12a, S2 identified the triangles formed by the diagonals drawn from one corner of the polygons. Besides, the student also showed that the triangle has no diagonal. Therefore, the student got three points. When we analyzed S2's answer, s/he explained exactly what the number n is, what (n-2) constitutes, and why 180o is written. Therefore, three points were given.

Findings obtained from the answers of the students with the lowest and highest scores

In this section, the answers of S6, who scored zero points in total in proof and justification, and S2, who scored nine points, to three questions were analyzed together. In this way, the answers of students at both levels were compared.

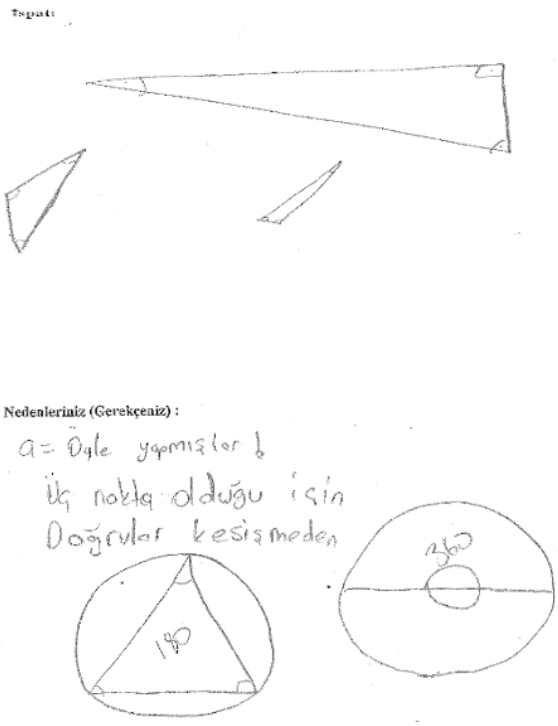
Figure 13.
S2 and S6's proofs and justifications about the sum of the measures of the triangle's interior angles
Sum of measures of a triangle's interior angles
S2
Writing proof



Justification
Transcription: A quadrilateral has 4 and the diagonal drawn from one vertex divides the shape into two triangles. Since the sum of the interior angles of the triangle is 180, we get 360 from 180.2

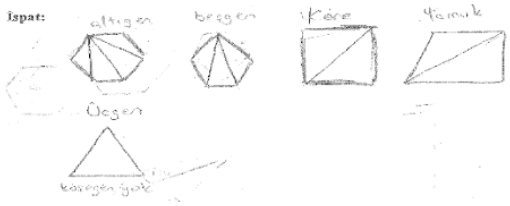
As can be seen from Figure 13, S2 did the proof and generalized because s/he wrote any quadrilateral in his/her justification and therefore s/he got full points. However, S6 tried to do the proof by taking different triangles. Similarly, s/he tried to justify, but he could neither provide complete proof nor justification. There is a big difference between his/her justification and his/her proof. Accordingly, the student received zero points.

S6
Writing proof



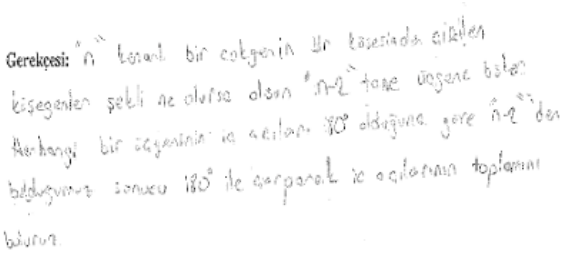
Justification
Transcription: Because they connect the four points without overlapping. because they found it that way

Figure 14.
S2 and S6's proofs and justifications about the sum of the measures of any quadrilateral's interior angles
Sum of the measures of any quadrilateral's interior angles
S2
Writing proof

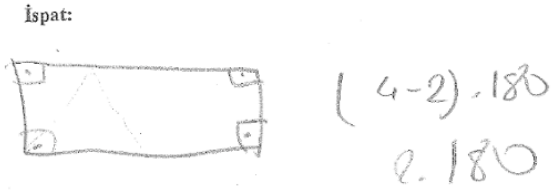


Transcription: Diagonals from one vertex of an n-sided polygon divide the shape into n-2 triangles. Since the interior angles of any triangle are 180, we can find the sum of the interior angles by multiplying n-2 by 180.

Justification

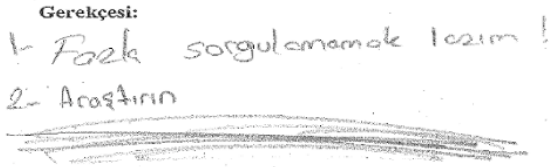


S6
Writing proof



Transcription: No need to questioning so much! Research it.

Justification



As can be seen from Figure 14, S2 showed the number of triangles drawn from a corner of different polygons in his/her proof. In his justification, s/he generalized this and proved the sum of the measures of the interior angles of a polygon with n sides. It was observed that S6 accepted the information without questioning. When the answers of S2 and S6 to three questions were analyzed, it was seen that S2 generalized simple to complex in the light of cause and effect relationship. S6, on the other hand, expressed it as an acceptance without questioning and questioning the cause.

Discussion and Conclusion

In the present study, 8th grade students' proof writing and justification skills in the context of the sum of polygon's angles were analyzed. In light of the findings, it was observed that the levels of justification skills were the same or one level below the levels of proof writing. There is only one student whose both proof writing and justification levels are "very good". This finding reveals the close relationship between proof writing and justification skills. It can be stated that the ability to justify is important for the development of proof writing skills. In addition, proof writing and justification skills are closely related to academic success.

In this study, the relationship between academic achievement and proof writing and justification skills was not examined. Indeed, Senk (1989) stated that the success of writing geometry proofs is positively related to the success of the geometry course. Özmusul (2018) stated that students with high achievement levels have high justification skills. Çalışkan (2012), on the other hand, stated that there is a positive correlation between the mathematics achievement of 8th grade students and their ability to prove. The findings of the Hatisaru (2020) study stated that the students perceive the mathematics classrooms as the teacher being at the center of education and training, explaining the subject and solving routine problems. Really teachers, curriculum, textbooks, etc. play an important role in the development of proof writing and justification skills. Polat (2015) analyzed the tasks requiring explanation and justification in two books determined by the Ministry of National Education for use in 7th grade mathematics teaching. The results of the study showed that tasks requiring explanation and justification are not included in mathematics textbooks.

Nine students were "unsuccessful" in writing proofs, four students were "moderately successful", three students were "very successful" and one student scored full points. The results of the study are consistent with the results of the studies stating that especially 8th grade students have difficulties with proof. Indeed, Albayrak Bahtiyari (2010) found that 8th grade students have deficiencies in both proof and reasoning. Similarly, Zaimoğlu (2012) found that 8th grade students could not fully comprehend the methods and techniques of proof. According to the results obtained from the justification skill levels, 11 students were unsuccessful, four students were "moderately successful", and one student was very successful with full points. This is consistent with the results of Özmusul (2018) that 7th grade students' complete and convincing justification skills are low, and Arslan (2007) that 6th, 7th, and 8th grade students' justification levels are low.

Another important result is that while the level of proof writing and justification level of eleven students were the same, the levels of six students were different. Except for one of these six students, the proof writing levels of the others were higher than the justification levels. If these two results are combined, the proof writing levels of the students participating in the study are either the same or higher than their justification levels. In this study, although it is related to the proof and justification of the sum of the measures of any polygon's interior angles, students are expected to make generalizations for any polygon based on triangle and quadrilateral. In other words, the proof of the sum of the measures of the interior angles of polygons is also related to generalization skills. The results of the study can also be explained by generalization skills. Indeed, there are studies in the literature that emphasize the relationship between generalization and justification skills (Ellis, 2007; Lannin, 2005; Radford, 1996). Ellis (2007) stated that justification affects a student's generalization ability. Similarly, Radford (1996) stated that justification is the process that supports generalization. Akkan, Öztürk, and Akkan (2017) stated in their study that pre-service teachers who generalize patterns correctly provide more justification than pre-service teachers who try to generalize or make no attempt. Yackel (2001) mentioned that providing justification and explanation has a positive effect on the mathematical norm in the classroom. Therefore, studies on justification as well as proof are important. As a result, the findings of this study are in parallel with the studies in the literature on both proof and justification.

Recommendations

The present study was carried out with 16 students who volunteered and participated in the face-to-face teaching process on the day the data would be collected due to the pandemic. This is the limitation of the study. Working with more students can be done.

In this study, proof writing and justification skills are discussed. The literature emphasizes that academic success is also important. Therefore, in future studies, metacognitive levels, academic achievement, geometric thinking levels, justification, and proof writing skills can be examined in detail.

As can be seen from the results of the study, students' proof writing and justification skills are not at the desired level. In this direction, activities aimed at improving students' justification skills can be included in teachers' lessons, textbooks, and curricula. In this way, classroom environments can be created where students can make inquiries, communicate, justify, and share their ideas easily. In addition, this study was conducted with a limited number of participants due to the pandemic. Future studies can be conducted with more participants.

Note: 1. This study was produced from the master thesis prepared by the second author under the supervision of the first author.

2. Within the scope of the research, ethics committee approval was obtained from the ethics committee of Sivas Cumhuriyet University with the decision dated 21.01.2021 and numbered 2021/26.

References

Akkan, Öztürk & Akkan (2017). Generalization processes of elementary mathematics teacher candidates: strategies and justifications. *Turkish Journal of Computer and Mathematics Education .8(3),513-550.*

Akkuş, R. (2019). Change in the level of justification in problem solving over time. *Kastamonu Education Journal, 27*(4), 1481-1494. doi:10.24106/kefdergi.3050

Albayrak Bahtiyari, Ö. (2010). *Awareness of proof and reasoning concepts and their importance in 8th grade mathematics teaching*. Unpublished Master Thesis. Ataturk University, Erzurum.

Alex, J. K. & Mammen, K. J. (2012). A survey of South African grade 10 learners' geometric thinking levels in terms of the van Hiele theory. *Anthropologist*, 14(2), 123-129.

Almeida, D. (2000). A survey of mathematics undergraduates' interaction with proof: some implications form mathematics education. *International Journal of Mathematical Education in Science and Technology*, *31*(6), 869-890. https://doi.org/10.1080/00207390050203360

Arslan, Ç. (2007). *Development of reasoning and proof thinking in primary school students*. Unpublished Doctoral Thesis, Uludağ University, Bursa.

Bayazıt, İ. (2017). Examining the importance of proof and teacher competencies in proof. *International Periodical for the Languages, Literature and History of Turkish or Turkic*, 12(14), 19-40.

Bell, A. W. (1976). A study of pupil's proof explanations in mathematical situations. *Educational Studies in Mathematics* 7(1), 23–40.

Çalışkan, Ç. (2012). Examination of the processes of processes of highly talent 8th grade students, Unpublished Doctoral dissertation, Bursa Uludag University.

Coşkun, F. (2009). The relationship between secondary school students' Van Hiele geometry comprehension levels and proof writing skills, Unpublished master's thesis. Karadeniz Technical University. Institute of Science and Technology, Trabzon.

Daguplo, M. (2014). How well do you write proof? Characterizing students proof-writing skill vis-à-vis van Hiele's model in geometrical proving. *Journal of Educational and Human Resource Development (JEHRD)*, *2*, 104-114.

Dimakos, G., Nikoloudakis, E., Ferentinos, S., & Choustoulakis, E. (2007). Developing a proof-writing tool for novice lyceum geometry students. *The Teaching of Mathematics*, *X*(2), 87-106.

Driscoll, M., Wing DiMatteo, R., Nikula, J., & Egan, M. (2007). Fostering geometric thinking: A guide for teachers, Grades 5-10. *Porsmouth,* NH: Heinemann.

Ellis, A. B. (2007). Connections between generalizing and justifying: Students' reasoning with linear relationships. *Journal for Research in Mathematics Education, 38*(3), 194–229.

Güner, P., & Topan, B. (2016). Prospective Elementary Mathematics Teachers' Abilities of Using Geometric Proofs in Teaching of Triangle. *Necatibey Eğitim Fakültesi Elektronik Fen ve Matematik Eğitimi Dergisi (EFMED)*, *10*(2), 210-242.

Goetting, M. M. (1995). *The college student's understanding of mathematical proof*. Doctoral dissertation. University of Maryland, College Park.

Hadas, N., Hershkowitz, R., & Schwarz B. B. (2000). The role of contradiction and uncertainty in promoting the need to prove in dynamic geometry environments. *Educational Studies in Mathematics,* 44, 127-150.

Harel, G., & Sowder, L. (1998). Students' proof schemes: Results from exploratory studies. In A. Schoenfeld, J. Kaput & E. Dubinsky (Eds.), *Research in collegiate mathematics education III* (pp. 234–283). Providence, RI: American Mathematical Society.

Hatisaru, V. (2020). School students' depictions of mathematics teaching and learning practices. *International Electronic Journal of Elementary Education*, *13*(2), 199-214.

Jones, K. (2000). The student experience of mathematical proof at university level. *International Journal of Mathematical Education in Science and Technology*, *31*(1), 53-60. https://doi.org/10.1080/002073900287381

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

Marrades R. & Gutierrez A. (2000). Proofs produced by secondary school students learning geometry in a dynamic computer environment. *Educational Studies in Mathematics* 44, 87–125. Netherlands, Kluwer Academic Publishers.

Mason, J., Burton, L., & Stacey, K. (2010). Thinking Mathematically. New York: Prentice Hall.

Milli Eğitim bakanlığı [MEB] (2020). *Elementary mathematics lesson 6-8th grades curriculum. Ankara: MEB.*

Milli Eğitim bakanlığı [MEB] (2014). Elementary Mathematics 7 Textbook. Ozyurt Printing. Ankara.

Miles, M. B. & Huberman, A. M. (1994). *Qualitative Data Analysis: An Expanded Sourcebook*, Sage.

National Council of Teachers of Mathematics NCTM (1989). *Principles and standards for school mathematics, Reston*/VA: National council of Teachers of Mathematics.

Polat, M. (2015). Examination of the tasks requiring explanation and justification in the elementary 7th grade mathematics courses and workbooks by learning areas. Unpublished Master's Thesis. Gaziantep University Institute of Educational Sciences.

Radford, L. (1996). Some reflections on teaching algebra through generalization. In L. Lee (Ed.), *Approaches to algebra: Perspectives for research and teaching* (107–111). Dordrecht, The Netherlands: Kluwer.

Ross, K. (1998). *Doing and proving: The place of algorithms and proof in school mathematics.* American Mathematical Monthly, 3, 252-255.

Senk, S.L. (1983). *Proof writing achievement and van hiele levels among secondary school geometry students*, Unpublished Master's Thesis, The University of Chicago.

Senk, S.L. (1985). How well do students write geometry proofs? *Mathematics Teacher*, Syracuse University.

Senk, S. L. (1989). Van Hiele levels and achievement in writing geometry proofs. *Journal for research in mathematics education*, *20*(3), 309-321.

Şen, C., & Güler, G. (2022). Examining proof-writing skills of pre-service mathematics teachers' in geometric proofs: van Hiele Model. *Journal of Kırşehir Education Faculty,* 23, 128-176.

Sowder, L., & Harel, G. (1998). Types of students' justifications. *The mathematics teacher*, *91*(8), 670-675

Stylianides, A. J., Stylianides, G. J., & Philippou, G. N. (2004). Undergraduate students' understanding of the contraposition equivalence rule in symbolic and verbal contexts. *Educational Studies in Mathematics*, *55*(1), 133-162. https://doi.org/10.1023/B:EDUC.0000017671.47700.0b

Stylianides, G. J., Stylianides, A. J., & Philippou, G. N. (2007). Preservice teachers' knowledge of proof by mathematical induction. *Journal of Mathematics Teacher Education*, *10*(3), 145-166. https://doi.org/10.1007/s10857-007-9034-z

Tanışlı, D., Yavuzsoy Köse, N., & Camci, F. (2017). Generalization and verification knowledge of prospective mathematics teachers in the context of patterns *Journal of Qualitative Research in Education, 5(3),* 195-222.

Özmusul, B. (2018). *Examination of secondary school 7th grade students' levels of mathematical justification skills.* Unpublished Master Thesis, *Gaziantep University, Gaziantep.*

Yackel, E. (2001). *Explanation, justification and argumentation in mathematic classrooms.* Proceedings of the Conference of the International Group for the Psychology of Mathematics Education (12-17 July, 2001). 1(4), Utrecht, Netherlands.

Yıldırım, A., & Şimşek, H. (2016). *Qualitative research methods in social sciences (10th ed.). Ankara: Seçkin Publishing.*

Zaimoğlu, Ş. (2012*). Geometric proof processes and tendencies of 8th grade students.* Unpublished Master's Thesis, Kastamonu University Institute of Science and Technology, Kastamonu.

iejee

International Electronic Journal of Elementary Education
September 2023, Volume 16, Issue 1, 31-42

Effectiveness of Virtual Tours to Archaeological Sites in Al-Ahsa in Developing Historical Concepts among Kindergarten Children

Asma Margeni Hussien Ali[a,*], Ftimah Sabah Abdulaziz Alowais[b]

Received : 12 March 2023
Revised : 19 July 2023
Accepted : 29 September
DOI : 10.26822/iejee.2023.312

[a,*] **Corresponding Author:** Asma Margeni Hussien Ali, Department of Kindergarten, Faculty of Education, King Faisal University, Saudi Arabia.
E-mail: ahussien@kfu.edu.sa
ORCID: https://orcid.org/0000-0002-6971-9772

[b] Ftimah Sabah abdulaziz Alowais, Department of Kindergarten, Faculty of Education, King Faisal University, Saudi Arabia.
E-mail: 222400502@student.kfu.edu.sa
ORCID: https://orcid.org/0000-0002-9796-6623

Abstract

The study aimed to reveal the effectiveness of virtual tours to the archaeological sites in Al-Ahsa in developing historical concepts among kindergarten children and to identify the differences between boys and girls in acquiring historical concepts. To achieve these objectives, a quasi-experimental approach based on one group design was used. The sample of the study consisted of (30) boys and girls between the ages of (5-6) years in the third level in the kindergarten attached to the third elementary school for early childhood in Hofuf in Al-Ahsa Governorate. The pictured historical concepts test was applied after calculating the coefficients of validity and reliability. The virtual tours to the archaeological sites in Al-Ahsa included "Al-Qarah Mountain, Suq Al-Qaysariyya, Ibrahim Palace, Jawatha Mosque, Al-Amiriya School, Al-Bay'ah House, Al-Uqair Port, and Al-Asfaar Lake". The results showed statistically significant differences at (0.05) between the means of the scores of the experimental group in the pre-and post-applications in the pictured test of historical concepts in favor of the post-application. Also, there was an effect size of virtual tours to archaeological sites on the development of historical concepts among kindergarten children. In addition, the results showed no statistically significant differences at (0.05) between the means of the scores of boys and girls in the historical concepts test pictured in the post application. The study recommended the inclusion of virtual tours and historical concepts in the self-learning curriculum for kindergarten.

Keywords:

Virtual Tours, Archaeological Sites, Historical Concepts, Kindergarten

Introduction

The current world is characterized by rapid growth in all fields and an information revolution in the field of educational technology. Electronic technologies developed, and modern educational patterns based on electronic technologies spread because of its many advantages represented in jumping over the barriers of time and space, overcoming problems related to traditional education, raising the level of independence of the learner, and motivating him to play positive educational roles.

The advent of the Internet contributed to the speed of access to knowledge and keeping pace with technological developments. It has become imperative for every educational institution to keep pace with these rapid developments. This is done by integrating technology into the educational process in a way that contributes to stimulating the child's motivation to learn and enhancing his senses. It also contributes to understanding abstract concepts difficult to acquire at this stage to achieve the desired educational outcomes (Al-Subaihi, 2019).Virtual tours through the Internet simulate the real environment and are considered a suitable means for the teacher and the learner. Through this, the learner acquires experiences and knowledge that may be difficult to achieve through traditional tours. They also make it possible to show tours to hard-to-reach places; either because of distance, like the South Pole and the planets, or the danger of reaching it, such as learning about the occurrence of volcanoes. They, in addition, develop the self-learning, research, and exploration skills of the learner. Furthermore, they can be done at any time and any place, and they can be performed more than once. In turn, this saves effort, time, and money. Further, they also contribute to providing a safe, interesting, fun, and attractive environment for learning and provide children with an environment similar to the real environment, which gives it an atmosphere of realism (Hassan, 2011). Historical concepts with human interaction with time and space and the process of recording and monitoring the historical facts and achievements of the ancestors and transmitting them to the current or future generations are considered of high priority. This helps in preserving the cultural and human heritage and transmitting it to future generations, following the example of the ancestors, and developing feelings of loyalty and belonging to the homeland (Badir, 2021).

Historical concepts are the cumulative nature of diverse and multiple knowledge. They make history and its facts, information, and concepts meaningful. History helps children understand how others live, and how they go about their lives spontaneously. It helps children feel more self-respect and learn to share ideas and roles, and how to get along with others. It also brings them up in the confines of civilization with its historical, cultural, scientific, and artistic depth. It is considered one of the most important entrances to education for the child, and it provides opportunities for him to know his heritage and his past and to link him to his present and future (Hijazi, 2020). The historical records in Al-Ahsa Governorate in the Eastern Province of the Kingdom of Saudi Arabia, from the pre-Islamic period to the history of the modern Saudi state, include a list of important evidence from these periods. This historical evidence includes cultural heritage resources, mosques, palaces, military fortresses, museums, schools, wells, etc. Also, there are naturally flowing hot and cold water springs. It is, in addition, famous for Limestone Mountains with caves and hollows. Accordingly, heritage resources and archaeological monuments played a positive role in the local economy by increasing the volume of internal and external tourism (Najm, 2019). Hence, giving the child the history of his country during his first years of education will have a lasting impact on him and affects a visible impact on his future and his life. The study of history broadens one's horizons, highlights the relationship between results and causes, stimulates thought, helps one face new situations, and predicts what will happen next (Badawi, 2006).

The Kingdom of Saudi Arabia has contributed to preserving the cultural heritage throughout history and passing it on to future generations. Educational tours are one of the important and appropriate strategies for children's developmental characteristics. They are a successful educational tool that arouses their interest in learning through the interesting and varied materials it offers. Despite the importance of educational field tours, they face many challenges and difficulties. The most important is the fear of teachers and parents of security and safety measures during the tour, or the lack of financial support for the implementation of such tours (Al-Subaihi, 2019). The process of acquiring the concept in the child is based on sensory perception and observing the surrounding things, people, events, generalization, and discrimination (Botros 2004). At the same time, historical concepts have a close relationship with chronological order. They refer to things or past events that a person has done in a past period, so it is difficult for children to understand them. Children suffer as the concept deepens, especially if it is related to time (Abdel-Wahhab et al., 2019).

Historical concepts are among the concepts that are difficult to present to the kindergarten child traditionally. Endocott's (2005) study aimed to search for some solutions to facilitate the acquisition of historical concepts for children due to the difficulty of the historical concept. The study concluded that children suffer as the concept deepens if it is related to time. Ismail's (2012) study measured the effectiveness of a multimedia computer program in imparting some historical concepts to kindergarten children. Muhammad (2015) also found the effectiveness of an activities program based on a simulation strategy in all its educational activities to develop some historical concepts among kindergarten children.

Based on the foregoing, the importance of learning and teaching historical concepts and archaeological sites has been demonstrated using organized programs with clear objectives. These programs require educational attitudes and strategies that are appropriate and beloved by the kindergarten child,

which enhance his motivation to learn, comprehend, and consolidate abstract historical concepts with ease. Given the difficulties that prevent kindergartens from implementing organized field tours, hence the need to use virtual tours arises to develop historical concepts and measure their effectiveness in order to benefit from and implement them in the educational field in kindergartens. Therefore, the current study, which is based on virtual tours of the archaeological sites in Al-Ahsa, aims to develop historical concepts among kindergarten children. This is done using panoramic virtual tours, whose content is presented in the form of 3D images for their suitability for this specific age group. They are also characterized by interesting elements that increase the effectiveness of the educational situation, ease of control and roaming freely within it and opportunities for searching and discovering aspects of the journey without leaving the classroom, through the application of (Google Earth). The study is expected to contribute to presenting a developed educational reality through traditional educational tours with a scientific professional vision through virtual tours. Also, it will show the procedural steps of virtual tours for kindergarten teachers in order to enable them to develop their performance and activate their use of virtual tours. In addition, it will provide guiding standards that may benefit virtual tour designers. Finally, it draws the attention of experts and curriculum developers in the field of childhood to the importance of integrating technology into the learning and teaching process. The problem of the study was reformulated in the following research questions:

1. What is the effectiveness of virtual tours to archaeological sites in Al-Ahsa in developing historical concepts among kindergarten children?

2. Are there any statistically significant differences at (0.05) between the mean scores of boys and girl in the historical concepts pictured test in the post-application?

Theoretical Framework

The theoretical framework deals with a description of virtual tours and historical concepts.

Virtual Tours

Virtual tours are an exciting component of the learning environment. Technology has helped in the development of educational means in an unprecedented manner and made its use something indispensable. The use of technology is no longer an unavoidable option in educational settings because it made a fundamental change in people's lives and their work. It has become an integral part of individuals' personal and practical lives (Faisal, 2014). Virtual tours contribute to enriching the educational process and motivating children's performance towards further excellence and progress and providing them with scientific and information outcomes that help them understand school subjects or public life. Tours enhance children's knowledge by linking them to their study units. In addition, they work on developing their skills and building their capacities by learning about global history, human history, and global technological developments within various virtual resources. Furthermore, they promote the values of tolerance and positive global citizenship among children by providing virtual sources and destinations such as the various historical, cultural, and religious sites of different peoples. They open the child's perceptions and develop his life and scientific and behavioral skills (Halawa, 2020). Piaget believes that education is not easy, as it needs planning and knowledge of the concepts that the learner can study at a certain age. Therefore, it is necessary to prepare and define the activities that the learner can carry out and to provide the opportunity to discover the information by himself and to focus in the education process on experimentation and exploration, not on indoctrination and memorization (Zaitoun, 2002). Therefore, virtual tours emerged as an alternative to actual field tours in order to enhance field work and enable educated children to solve these limitations, through modern technological education applications. The field of virtual tours is constantly being improved, developed, and renewed, as it is possible to create very advanced virtual tours capable of replacing real field tours (Caliskan, 2011).

Virtual tours are one of the modern technological innovations that appeared through the Internet. The first use of a virtual tour was in 1994 by Queen Elizabeth II, when the visitor center was officially opened and was called a "Virtual Tour". It was conducted through virtual reality. Virtual tours have become an effective means used via the Internet to display some archaeological and tourist sites and museums and use them in education. Through this, maximum benefit can be achieved in classrooms because they help learners to understand the different curricula (Hassan, 2010). When planning educational virtual tours, their objectives must be clear and related to the content of the curriculum activities. Parents should also participate in them and cooperate whenever possible. Tours motivate children to think about the environment and express their experiences and benefit from them. They also provide suitable opportunities for children to acquire facts and concepts related to the natural, human, and social phenomena surrounding them. They are an opportunity to identify, change, and remove the feeling of boredom from the children’s souls by moving to other places different from their kindergarten and homes, opening them the opportunity to go out and express themselves (Sharif, 2015). Virtual tours

are an interactive online environment that simulates any location using panoramic images. They include multiple media such as text, images, sound effects, and audio clips, and allow the learner to obtain the information he wants and contributes to the development of the skills he needs (Khamis, 2016). They are a technological innovation that allows the learner to learn in a safe and attractive virtual environment anywhere and at any time, by activating several technologies to achieve planned educational goals (Al-Sobhi, 2019).

Types of Virtual Tours

Al-Meligy (2020) explained the types of virtual tours. Frist, text-based virtual tours are tours that depend on a detailed presentation of the tour through the use of written texts. This type of virtual tour is the simplest and least expensive. It does not use optical tools. Second, sound-based virtual tours rely on audio clips, by displaying content that includes sound effects such as the sound of sea waves or the sound of footsteps. Third, video-based virtual tours depend on the presence of video copies that are completely identical to the original copies of the tour. These tours feature audio and text feedback about the contents of the tour. However, one of the disadvantages of this type of tour is the high cost of production as well as the difficulty of continuously updating it. Fourth, panoramic virtual tours give learners a great sense of reality. They present their content in three-dimensional form and rely on the presence of a group of images that are linked together to form a 360-degree panorama. In this type of tour, accuracy is taken into account in the production and assembly of images, which must be of high quality. Fifth, three-dimensional virtual tours depend on a group of three-dimensional scenes, which are kinetic. These scenes allow the learner to control and interact with the elements of the journey, such as walking from one place to another, as if moving in a real environment simulating reality. Sixth, synchronous virtual tours are the most attractive and exciting types. They are a combination of video-based and panoramic flight. This type of tour allows the learner to wander in a three-dimensional environment simulating the real environment, using tour tools and is considered one of the most expensive types of virtual tours. No matter how many types of tours are, activating them in education requires defining the purpose of their use and linking them to the content and method of teaching. The male or female teacher may depend on one of the types according to the capabilities available to him (Al-Sobhi, 2019). According to the objectives of the current study, the study relied upon panoramic virtual tours, the content of which was presented in a three-dimensional form. It relies on the presence of a set of images that are linked together to form a 360-degree panorama, due to its suitability for this specific age group of the research sample. Also, they provide ease of use and navigation within the software at a low cost.

Characteristics of Virtual Tours

Virtual tours provide a great deal of interaction between the environment and the user. They also provide realistic ways to access and explore the different components of the environment. They have a lower material cost compared to real tours. They are easily accessible since they can be published on the Internet. In addition, they are subject to continuous modification and updating due to their flexibility. They are characterized by the sharing of learning resources and the provision of products to individual participants. The user can move within the environment without any restrictions. Finally, they allow visual information and data to be visualized graphically so that it appears real (Falih & Harith, 2021). The advantages of virtual field tours are: integrating various types of data in immediately available ways, providing images from a variety of different scales, and displaying invisible data such as corn, geochemistry, etc. Also, they are useful for offering tours to inaccessible areas and enable the display of extensive field tours with a great variety of landform diversity. In addition, they enhance and expand students' experience and enable flexibility of access through time and space. Furthermore, they provide a repeatable experience and can be used to reinforce concepts in the classroom (Caliskan, 2011). Khamis (2016) also summarized the factors that require the use of virtual tours, including the spread of Internet services and their low cost, low prices of computers and the spread of laptops, directing learners to internet sites and expanding their use, and what real tours face in terms of obstacles that prevent their employment in light of security and economic conditions in the communities, including lack of administrative support., increasing expenses, and lack of sufficient time.

Al-Meligy (2020) revealed that using virtual tours software has high effectiveness in developing archaeological awareness and strengthening the values of national belonging among children in the research group. Zain El Din (2018) also demonstrated the effectiveness of 3D virtual tours in providing kindergarten children with basic geographic concepts in the developed curriculum for kindergarten and came up with a list of technical and educational standards for virtual tours. In addition, Hussein (2020) showed the impact of providing performance support in virtual learning tours to develop the life skills of kindergarten children. Moreover, Badir and Khamis (2020) indicated the impact of the museum trips program on tourism development for children in the early childhood stage in the Kingdom of Saudi Arabia and tourism awareness. Finally, Faleh and Harith (2021) concluded that there is a relationship between

panoramic virtual trips and the development of the concept of color among kindergarten children.

Historical Concepts and Children

Historical concepts are one of the educational fields concerned with the study of man, in terms of his growth, way of thinking, behavior, and the natural surroundings in which he lives through his history and its intellectual, social, economic, and political differences. History represents a prominent place among the branches of human knowledge. It is the totality of what happened previously in terms of historical events, which discovers the roots of the present human problems and the results that led to them. It is also a children's window to look at the past of their country and their leaders, and what civilization went through in light of the experiences of the former (Budair & Khamis, 2020). Some educators stress the importance of starting to learn historical concepts at an early age, on the condition that an appropriate traditional environment is provided, and that various and different activities are made available, in a manner commensurate with the child's nature, trends, and inclinations. Therefore, the acquired concepts of this stage must be included and integrated into the activities that the child interacts with. We may find that the society in which the child lives is a living, ever-evolving, and changing society. Therefore, the child must feel the impact of the past in the present, and what is expected of the impact of the present in the future (Al-Hayla, 2001).

The importance of learning historical concepts for the kindergarten child is concentrated in the following: identifying the natural and cosmic phenomena that children see and seek to try to explain them, developing the child socially, emotionally, and intellectually through various events, and making him more efficient in dealing with his environment. They allow summarizing surrounding information about the child, which makes it easier to store and retrieve it from memory and deal with it in different situations of life, or educational situations, so the child memorizes as much vocabulary as possible that helps him retrieve it and deal with it flexibly at the time of need. In addition, every child's learning is based on a sense of time and place that helps to learn historical concepts in their later learning and contributes to organizing the child's mental experiences. Historical concepts are a starting point or a guiding point, and a path for the various activities of the child, as they may be a result of the educational process, or be a constructive means for lessons or other concepts. Moreover, the teaching of historical concepts is an important factor in correcting incorrect or abstract concepts alike. Likely, historical concepts enable the transfer of learned scientific experiences to new situations without the need to learn to solve each problem separately, so it is sufficient to generalize to all similar situations. Furthermore, historical concepts elevate the child to a level where he can extract new ideas, laws, and explanations for a number of phenomena and events, which is the level at which creativity and innovation of all kinds appear. They simplify and organize the environment surrounding children and facilitate children's communication with each other. Finally, they help to identify the events and facts, the date of their occurrence, and the accompanying changes and events.

The development of historical concepts among kindergarten children aims to introduce children to the history of their country, prominent personalities, and important historical events and to encourage them to respect their historical heritage and belonging to their homeland. It also satisfies the motive of knowledge and curiosity of the child in knowing himself and understanding his relations with the past and his relations with other societies and their culture. It, in addition, develops the child's ability to understand historical events through various historical activities and teaches the child how to think within his capabilities and realize the relationships between knowledge, reason, and historical events. Moreover, it helps the child to be acquainted with the civilization of his country and to try to strive for the elevation and progress of his country, to be attached to it and his land, to take lessons and cues from the events of the past, and to benefit from its lessons in addressing the present and the future. Finally, it helps children to realize the great effects that Islam has made in the history of humankind in general and to recognize the importance of historical places that attract tourists to their homeland (Bawazir & Qurban, 201; Chick, 2006; Ibrahim & Muhammad, 2016).

The foundations that the kindergarten teacher must take into account when teaching historical concepts are to start teaching tangible concepts, then abstract ones, then more abstract ones, and choose the lower educational levels, then the upper levels according to the children’s levels. Care must also be taken on the practical application when teaching historical concepts and setting a good example for children. In addition, when choosing scientific experiences, the nature of the learner himself must be taken into account in terms of his inclinations, interests, and curiosity. Humans and nature influence each other directly or indirectly. Finally, when developing plans and curricula, teachers must take into account the individual differences between children's cultural, religious, social, and mother tongues (Botros, 2004; Cohen, 2010; Dikmenli, 2014; Zahran et al., 2007). Ibrahim and Muhammad (2016) also demonstrated that the training program motivated kindergarten teachers to use the story for the kindergarten child to acquire some historical concepts in the Pharaonic

era after their knowledge of this era. Also, Al-Mounir et al.'s (2019) study revealed the development of some historical concepts among kindergarten children using a program based on the systemic approach. In addition, Badir (2021) demonstrated the effectiveness of using the virtual museum in developing historical concepts in early childhood.

Aim of the Study

The current study aims:

- to discover the effectiveness of virtual tours to archaeological sites in Al-Ahsa in developing historical concepts among kindergarten children.
- to detect differences between boys and girls in acquiring historical concepts.

Methodology

The quasi-experimental approach was used, given that it is appropriate for the nature of this study, achieving its objectives, and revealing the effectiveness of virtual tours to the archaeological sites in Al-Ahsa in developing historical concepts among kindergarten children. The study also relied on the design with one treatment group. Quasi-experimental studies evaluate interventions but that do not use randomization. Also, they demonstrate causality between an intervention and an outcome.

Population and Sample of the Study

The study population consisted of all kindergarten children in the third level for the first semester of the academic year 2022, in government kindergartens affiliated with early childhood schools in Al-Ahsa Governorate in the Eastern Province. The cluster sampling technique was applied to draw the study sample. One early childhood school (the Third Elementary for Early Childhood in Hofuf) was selected. The researchers chose (30) boys and girls to participate in the study based on their parents' approval. Those parents who agreed to allow their children's participation in the study completed two copies of the written informed consent form. They kept one copy and returned the other to the researchers. Also, the approval letter for conducting the research from the Ethical Approval Committee at the Deanship of Scientific Research, Kind Faisal University was obtained. They were recruited in the study in the first semester of 2022. It must be noted that authors did not have access to information that could identify individual participants during and after data collection. The participants are considered homogenous in terms of nationality (Saudi Arabia), culture, and economic status.

Instruments

The study tools consisted of a pictured test of historical concepts and a program based on virtual tours to develop historical concepts. After reviewing the theoretical framework and previous studies related to historical concepts (Al-Meligy, 2020; Bdeir & Khamis. 2020), the pictured historical concepts test was built. The test aimed to measure the extent to which kindergarten children acquired historical concepts. It has been prepared according to the following steps:

- The initial image was prepared for the pictured historical concepts test, which consisted of (16) paragraphs distributed over eight main dimensions. The first dimension includes Jabal Al-Qarah, the second Suq Al-Qaisariya, the third Ibrahim's Palace, the fourth Jawatha Mosque, the fifth Al-Uqair Port, the sixth Al-Amiriya School, the seventh Al-Bay'ah House, and the eighth Al-Asfaar Lake. Each dimension contains two sections; the first section measures the child's ability to know the shape of the historical landmark, and the second dimension measures the child's recognition of the area in which the historical landmark is located. The instructions were drafted, and a correction key was put in place.

- The initial picture of the test was shown to a jury of experts.

- The test was applied to the survey sample.

- The test in its final version consists of two parts: the first part is concerned with the child's basic data: name, gender, date of birth, date of application of the test, and its instructions. The second part includes questions distributed over the main criteria of historical concepts. Under each question, there are three pictures from which the child can choose the one that answers the correct answer.

- The test was corrected by giving one point if the answer was correct, zero point if the answer was wrong, and the total score was 16 degrees.

Test Validity

To verify the face validity of the test, it was presented to a number of (7) experts who are specialists in early childhood, educational techniques, and teaching curricula and methods.

Their comments and suggestions were taken into account in developing the test. Also, the test was piloted on a sample of (10) children. Then, the Pearson correlation coefficient was calculated between the child's score on the question and the total score on the test. Table 1 shows the results.

Table 1.
Pearson Correlation Coefficients between the Children's Score on the Question and the Total Score of the Pictured Test

No. of question	Pearson correlation coefficient	Sig.	No. of question	Pearson correlation coefficient	Sig.
1	.728*	.017	9	.838**	.002
2	.692*	.027	10	.692*	.027
3	.728*	.017	11	.633*	.049
4	.633*	.049	12	.768**	.009
5	.692*	.027	13	.654*	.040
6	.633*	.049	14	.838**	.002
7	.654*	.040	15	.768**	.009
8	.753*	.012	16	.809**	.005

**sig. at (0.01), * at (0.05).

Table 1 shows that the Pearson correlation coefficients between the scores of the test questions and the total score of the test were statistically significant at (0.01) and (0.05). Pearson's correlation coefficients ranged between questions with the total score of the test between (0.633* - 0.838**). This result confirms that the test has a high degree of validity.

Test Reliability

The reliability coefficient was calculated on the total score of the test using the Couder-Richardson equation -20. The test was applied to a survey sample, consisting of (10 children). The reliability coefficient was (0.93). The reliability coefficient was calculated through the split-half (Brown), and the reliability coefficient was (0.87). These values are high-reliability coefficients that are suitable for the study.

Virtual Tours-based Program

The program was built by looking at several sources represented in books, references, studies, and previous research that focused on virtual tours, sites, and historical concepts. The program consists of eight sessions that contain some historical concepts in Al-Ahsa Governorate in the Kingdom of Saudi Arabia, namely, Al-Qarah Mountain, Suq Al-Qaisariya, Ibrahim Palace, Jawatha Mosque, Al-Amiriya School, Bait Al-Allegiance, Al-Uqair Port, and Al-Asfaar Lake. The program aims to acquaint children of the third level in kindergarten with the antiquities in their environment and urge them to discover more about what the Al-Ahsa Governorate contains preserved and valuable historical antiquities over time. The program was carried out using panoramic virtual tours presented in 3D. It depends on the presence of a set of images linked together to form a 360-degree panorama because it is suitable for kindergarten children, as it provides ease of navigation, use, and navigation within the software, and its low cost through the application (Google Earth). The program includes objectives, means, strategies, and evaluation methods.

The content of the program was presented in its initial form to (7) experts specialized in the field of early childhood, educational technologies, curricula, and teaching methods, in order to take their opinions on the appropriateness of the general objective and the detailed objectives and to amend or delete what they deem appropriate. The experts' remarks about modifying some of the session's elements and activities and amending the wording of some procedural objectives with the content of the sessions were taken into consideration. In its final version, the program consisted of (8) sessions. Table 2 shows a summary of the content of the program sessions, its objective to be achieved, the necessary tools and means, and the proposed time for each session.

Table 2.
Summary of the Program Content

The overall goal of the program	Developing historical concepts through virtual tours for kindergarten children.
Session topic	The first session: the virtual tour to Al-Qarah Mountain The second session: the virtual tour to Suq Al-Qaisariya The third session: the virtual journey to Ibrahim's palace Fourth Session: Virtual Journey to Jawatha Mosque Fifth session: the virtual tour to Al-Amiriya School Sixth session: Virtual tour to Al-Bay'ah House Seventh session: Virtual tour to Al-Uqair Port Eighth session: the virtual tour to Al-Asfaar Lake
Session time	30-35 minutes
Used means	A smart tablet, projector or (smart board) to display virtual tours.
Activity course	-The introductory part includes a pre-diagnostic assessment. -The main part includes a phased (structural) assessment. -Concluding part.
Summative assessment	-A final assessment that includes questions about the virtual tours offered to determine the extent to which the child has achieved the objectives of each session.

Data Analysis

SPSS software (v. 23) was used to analyze the data. A number of statistical methods were used:

- Pearson correlation coefficient to calculate the validity of the internal consistency of the test.

-Couder-Richardson equation 20 to calculate the reliability of the test.

- T-test for paired and independent samples to verify the first and second hypotheses.

- The size of the effect to find out the effectiveness of virtual tours to archaeological sites in Al-Ahsa in developing the historical concepts of the kindergarten child.

Results and Discussion

What is the effectiveness of virtual tours to archaeological sites in Al-Ahsa in developing historical concepts among kindergarten children?

To verify the validity of the first questions, the t-test was used for paired samples to show the significant differences between the means of the scores of the children of the experimental group in the pre-and post-applications of the pictured historical concepts test. Table 3 shows the results.

It can be seen from Table 3 that there were statistically significant differences at (0.05) between the means of the scores of the experimental group in the pre-and post-tests in all dimensions of the test, in favor of the post-test. The level of the effect size of the effectiveness of virtual tours to the archaeological sites in Al-Ahsa in developing historical concepts in all dimensions of the test was high. This result validates the first hypothesis. Therefore, virtual tours to the archaeological sites in Al-Ahsa were effective in developing historical concepts among kindergarten children. This result is because the virtual tours embody the real reality of the historical places with their reliance on visual and kinetic effects that draw the child's attention towards acquiring and developing historical concepts. In addition, the virtual tours were more interesting for the child to listen to historical concepts than the usual traditional methods. Therefore, virtual tours were effective in developing historical concepts among kindergarten children.

Table 3.

The T-Test of Paired Samples for the Differences between the Means of the Scores of the Children of the Experimental Group in the Pre-and Post-Applications of the Pictured Historical Concepts Test

Dimensions of the Historical Concepts Test	Group	No.	Means	Standard deviations	t	df	Sig.	Differences	Effect size	Level
Al-Qarah Mountain	Pre-test	30	.90	.845	6.656	29	.000	Post-test	1.21	High
	Post-test	30	1.93	.254						
Suq Al-Qay-sariyya	Pre-test	30	.93	.740	6.513	29	.000	Post-test	1.01	High
	Post-test	30	1.87	.346						
Ibrahim Pal-ace	Pre-test	30	.63	.718	7.526	29	.000	Post-test	1.38	High
	Post-test	30	1.87	.346						
Jawatha Mosque	Pre-test	30	.33	.479	14.355	29	.000	Post-test	2.62	High
	Post-test	30	1.83	.379						
Al-Uqair Port	Pre-test	30	.57	.679	11.195	29	.000	Post-test	2.04	High
	Post-test	30	1.93	.254						
Al-Amiriya School	Pre-test	30	.43	.626	9.633	29	.000	Post-test	1.76	High
	Post-test	30	1.77	.504						
Al-Bay'ah House	Pre-test	30	.53	.681	11.195	29	.000	Post-test	2.04	High
	Post-test	30	1.90	.305						
Al-Asfaar Lake	Pre-test	30	.57	.626	12.339	29	.000	Post-test	2.26	High
	Post-test	30	1.97	.183						
Total degree	Pre-test	30	4.90	2.023	24.954	29	.000	Post-test	4.56	High
	Post-test	30	15.03	1.520						

The result of this study agrees with that of Zain El Din's (2018) study, which demonstrated the effectiveness of 3D virtual tours in providing kindergarten children with basic geographic concepts in the developed curriculum for kindergarten at Port Fouad Experimental Language School in Port Said. It also converges with that of Hussein's (2020) study, whose results found an effect of the method of providing performance support in virtual learning tours to develop the life skills of kindergarten children. In addition, this result also intersects with that of Al-Meligy's (2020) study, which showed that using virtual tours software has great effectiveness in developing archaeological awareness of Egyptian antiquities and strengthening the values of national belonging among kindergarten children. Moreover, this result is consistent with the result of Falih and Harith's (2021) study, which found a relationship between panoramic virtual tours and the development of the concept of color in children from the governmental governorate of Riyadh. Furthermore, Budair and Khamis's (2020) study supports the current result. The results showed statistically significant differences between the averages of the post-test for each of the experimental and control groups in favor of the experimental group in the degree of awareness of kindergarten children in the Kingdom of Saudi Arabia. This result indicates the impact of the museum tours program on tourism development for early childhood children in the Kingdom of Saudi Arabia. Finally, the current result accords with that of Budair's (2021) study, which revealed the effectiveness of the virtual museum in developing historical concepts in early childhood.

Reporting and discussion of the results of the second questions: Are there any statistically significant differences at (0.05) between the mean scores of boys and girls in the historical concepts test pictured in the post-application.

To verify the validity of the second questions, the t-test was used for independent samples to show the significance of the differences between the average scores of males and females in the concepts test in the post-test of the historical concepts test. Table 4 shows the results.

Table 4.

T-test for Independent Samples of the Differences between the Average Scores of Boys and Girls in the Concepts Test in the Post-test of the Historical Concepts Test

Dimensions of the Historical Concepts Test	Gender	No.	Means	Standard deviations	t	df	Sig.
Al-Qarah Mountain	Boys	14	1.93	.267	.095	28	.925
	Girls	16	1.94	.250			
Suq Al-Qaysariyya	Boys	14	1.79	.426	1.209	28	.237
	Girls	16	1.94	.250			
Ibrahim Palace	Boys	14	1.79	.426	1.209	28	.237
	Girls	16	1.94	.250			
Jawatha Mosque	Boys	14	1.86	.363	.317	28	.754
	Girls	16	1.81	.403			
Al-Uqair Port	Boys	14	1.93	.267	.095	28	.925
	Girls	16	1.94	.250			
Al-Amiriya School	Boys	14	1.79	.426	.190	28	.850
	Girls	16	1.75	.577			
Al-Bay'ah House	Boys	14	1.86	.363	.714	28	.481
	Girls	16	1.94	.250			
Al-Asfaar Lake	Boys	14	2.00	.000	.933	28	.359
	Girls	16	1.94	.250			
Total degree	Boys	14	14.93	1.492	.348	28	.731
	Girls	16	15.13	1.586			

Table 4 shows no statistically significant differences at (0.05) between the mean scores of boys and girls in the historical concepts test pictured in the post-test. The level of significance on all dimensions and the total score was greater than (0.05). This result indicates that the effectiveness of virtual tours to the archaeological sites in Al-Ahsa has an effect of the same degree on the development of historical concepts among kindergarten children, both boys and girls.

This result is attributed to the fact that virtual tours have a positive impact on the development of historical concepts among children of both the boys and girls sexes. In addition, both boys and girls children have the same enthusiasm, motivation, and interaction. According to the cognitive theory, the gender factor does not affect the acquisition of knowledge in the kindergarten stage. The child goes through developmental changes that appear as a result of the experiences he is exposed to, whether he is boys or girls as their needs and interests are similar at this age. The result of this study agrees with that of Al-Mounir et al.'s (2019) study, which aimed to develop some historical concepts among kindergarten children using a program based on the systemic approach. The study used the pictured historical concepts test for kindergarten children, and the program was based on the systemic approach. The study showed no statistically significant differences between the mean scores of boys and girls in the concepts test in the post-application.

Conclusion

The study aimed to reveal the effectiveness of virtual tours to the archaeological sites in Al-Ahsa in developing historical concepts among kindergarten children and to identify the differences between boys and girls in acquiring historical concepts. The results showed statistically significant differences between the means of the scores of the experimental group in the pre-and post-applications in the pictured test of historical concepts in favor of the post-application. Also, there was an effect size of virtual tours to archaeological sites on the development of historical concepts among kindergarten children. In addition, the results showed no statistically significant differences between the means of the scores of boys and girls in the historical concepts test pictured in the post application. The study results implicate the importance of developing kindergarten children's historical concepts using virtual tours. Based on the results, the researcher recommends paying attention to the use of virtual tours in the kindergarten stage because of their acceptance and importance in achieving children's acquisition of historical concepts and educational goals and increasing the speed of learning among kindergarten children. Parents of children should also be invited to employ virtual tours during their children's education and to choose from virtual tours that are suitable for their children's ages. In addition, there is a need to train kindergarten teachers through training workshops and educational seminars on the use of virtual tours. Finally, the researcher suggests conducting a study on the use of virtual tours in the development of scientific and social concepts.

Acknowledgment

This work was supported by the Deanship of Scientific Research, Vice Presidency for Graduate Studies and Scientific Research, King Faisal University, Saudi Arabia [Project No. GRANT4,129].

References

Abdel-Wahhab, A. G., al-Mahlawi, G. M., & Radwan, T. R. (2019). The role of stories in developing some historical concepts for kindergarten children. *Journal of the Faculty of Specific Education for Educational and Specific Studies,* 9, 36-52.

Al-Hayla, M. M. (2001). *Educational and information technology*. University Book House, Al Ain, Emirates.

Al-Meligy, R. R. (2020). The effectiveness of using virtual tours to develop archaeological awareness and support the values of national belonging among kindergarten children. *Journal of Studies in Childhood and Education* - Assiut University, (14) 319-374. Retrieved from https://journals.ekb.eg/article_137786.html

Al-Mounir, R. A., Ali, N. A., & Ahmed, M. Y. (2019) .The effectiveness of a program based on the systemic approach in developing some historical concepts and deductive thinking among kindergarten children. The second international conference: *Building the fourth generation child in the light of Education Vision 2030*, Assiut: Assiut University. Kindergarten College, 427-450.

Al-Sobhi, N. S. (2019). *Employing virtual tours in the educational process*. New education. Tarawneh. Retrieved from https://jwadi.journals.ekb.eg/article_84873_c41c75a2cf9cbd4883ae66be9c34a6a9.pdf

Badawi, A. M. (2006). *The science of history, its usefulness, and its educational functions in our changing world between theory and practice*. Modern Book House.

Badir, K. M. (2021). The effectiveness of the virtual museum in developing historical concepts in early childhood. *Journal of Studies in Childhood and Education* - Assiut University, *16*(16), 1-26.

Badir, K. M., & Khamis, S. I. (2020). The impact of a museum tour program on developing tourism awareness for early childhood children in the Kingdom of Saudi Arabia. *Journal of the Faculty of Education*, Mansoura University, (112), 531-555. Retrieved from https://journals.ekb.eg/article_180776.html

Bawazir, S. A., & Qurban, N. A. (2011). *Developing historical and geographical concepts for kindergarten children*. Amman: Dar Al-Masirah for publishing and printing.

Botros, H. B. (2004). *Developing scientific concepts and skills for pre-school children*. Al Masirah House: Amman.

Çaliskan, Onur. (2011). Virtual field tours in education of earth and environmental sciences, 3rd World Conference on Educational Sciences, *Procedia Social and Behavioral Sciences ,15*, 3239–3243. https://www.sciencedirect.com/science/article/pii/S187704281100824X

Chick, K. A. (2006). Fostering student collaboration through the use of historical picture books. *The Social Studies, 97*(4), 152-157.

Cohen, L. E. (2010). Exploring Cultural Heritage in a Kindergarten Classroom. *NAEYC Journal, 3* (15), 72-77.

Dikmenli, Y. (2014). "Geographic literacy Perception Scale (GLPS): Validity and Reliability Study. *Mevlana International Journal of Education,* 4(1), 1-15.

Endocott, J. (2005). Its not all history now: connecting the past by weaving a threaded historical concept. *The social studies, 96*(5), 227-232.

Faisal, Y. (2014). *Developing three-dimensional virtual tours to develop some life skills for kindergarten children* [Unpublished master's thesis]. College education quality, Port Said University.

Falih, H., & Harith, S. (2021). Panoramic Virtual Tours and their Relationship to Develop the Concept of Color among Kindergarten Children. *PalArch's Journal of Archaeology of Egypt/ Egyptology, 18*(4), 1666-1684. Retrieved from https://www.archives.palarch.nl/index.php/jae/article/view/6554

Halawa, R. (October 14, 2020). *Virtual school tours enrich the educational process and motivate excellence*. Al-Bayan newspaper. Retrieved from https://www.albayan.ae/across-the-uae/education/2020-10-14-1.3985813

Hassan, R. A. (2010). *A suggested model for online virtual tours and its effectiveness in developing educational technology students' achievement and their attitudes towards it* [Unpublished master's thesis]. Faculty of Specific Education, Ain Shams University.

Hassan, R. A. (2011). Standards for building virtual tours online. Education Technology. *Studies and Research by the Arab Society for Educational Technology*, 125-152. Retrieved from http://search.shamaa.org/FullRecord?ID=91311

Heritage Authority. (D.T.). *Heritage Sites in the Eastern Province*. Ministry of Culture.

Hijazi, H. Y. A. (2020). The effectiveness of a program based on library activities on developing some geographical and historical concepts for kindergarten children. *Scientific Journal of Educational and Specific Studies and Research, College of Specific Education*, (11), 89-100.

Hussein, R. R. (2020). The effect of providing performance support in virtual tours on the development of life skills for kindergarten children. *Arabic Studies in Education and Psychology, 9* (118) 249-270. Retrieved from https://journals.ekb.eg/article_69657.html

Ibrahim, H. H., & Muhammad, W. M. (2016). A training program for kindergarten teachers based on the use of the story to acquire some historical concepts in the Pharaonic era for the kindergarten child. *Childhood and Education Journal,* (25), 93-170.

Ismail, A. (2012). *The effectiveness of an interactive multimedia computer program to acquire some historical concepts for kindergarten children* [Unpublished master's thesis]. Faculty of Kindergarten, Cairo University.

Khamis, M. K. (2016). The effectiveness of a proposed program based on virtual tours via the web in teaching geography to develop the dimensions of geographical culture among middle school students. *Arabic Studies in Education and Psychology 73*(73), 69-109. Retrieved from http://search.mandumah.com/Record/760952

Ministry of Culture (D. T). *Culture House*. Ministry of Culture.

Ministry of Culture (D. T). *House of Allegiance Museum*. Ministry of Culture.

Ministry of Culture. (D.T.) *Al-Ahsa: eyes, palm trees, and welcoming faces*. Ministry of Culture.

Najm, W. H. A. K. (2019). Designing a tourist advertising poster in Al-Ahsa Oasis in the Eastern Province of Saudi Arabia as an introduction to teaching decorative designs in the field of advertising. *Journal of Specific Education Research - Mansoura University*, (53), 104-136.

Sharif, M. A. (2015). *Social and religious education in kindergarten*. Amman: Dar Al-Masirah for publishing and printing.

Shehata, H., & Al-Najjar, Z. (2013). *Dictionary of educational and psychological terms* (4th ed.). The Egyptian Lebanese House.

Zahran, H., & Rushdi, T., & Taima, A. (2007). *Children's linguistic concepts, their foundations, skills, teaching, and evaluation*. Amman: Dar Al-Masirah for publishing and printing.

Zain El-Din, M. M. (1018). The effectiveness of 3D virtual tours in providing kindergarten children with basic geographic concepts in the developed curriculum. *Journal of the Ismaili College of Education*, (40). 129-192. Retrieved from https://jfes.journals.ekb.eg/article_87226_9a75e6fd666ea54dba5a55f385def9fb.pdf

Zaitoun, K. A. (2002). *Teaching science to understand*. Cairo: World of books for publishing, distribution and printing.

iejee

International Electronic Journal of Elementary Education
September 2023, Volume 16, Issue 1, 43-55

Influence of Advantages and Levels of Reflection of Podcasts On Communicative Competences

Fabiola Talavera-Mendoza[a,*], Héctor Exequiel Gamero Torres[b], Alicia Miguelina Vera Manchego[c], María Isabel Benites Gamio[d], Karen Yosio Mamani Monrroy[e]

Received : 5 May 2023
Revised : 4 August 2023
Accepted : 29 September 2023
DOI : 10.26822/iejee.2023.313

[a,*] **Corresponding Author:** Fabiola Talavera - Mendoza, Universidad Nacional de San Agustín de Arequipa.
E-mail: ftalaveram@unsa.edu.pe
ORCID: https://orcid.org/0000-0002-0008-5206

[b] Héctor Exequiel Gamero Torres, Universidad Nacional de San Agustín de Arequipa.
E-mail: hgamerot@unsa.edu.pe
ORCID: https://orcid.org/0000-0003-0537-4629

[c] Alicia Miguelina Vera Manchego, Universidad Nacional de San Agustín de Arequipa.
E-mail: averaman@unsa.edu.pe
ORCID: https://orcid.org/0000-0003-0253-0542

[d] María Isabel Benites Gamio, Universidad Nacional de San Agustín de Arequipa.
E-mail: mbenitesg@unsa.edu.pe
ORCID: https://orcid.org/0000-0002-3947-4583

[e] Karen Yosio Mamani Monrroy, Universidad Peruana Unión.
E-mail: karen.mamani@upeu.edu.pe
ORCID: https://orcid.org/0000-0002-4363-1772

Abstract

An instrument was validated through factorial structure to measure the advantages, levels of reflection, and communicative competencies of the podcast, as well as to establish the levels of reliability and consistency of the measuring instrument. Likewise, confirmatory analysis was carried out to validate the proposed structural model and establish the relationships of the factors using the PLS-SEM methodology, Partial Least Squares Structural Equation Modeling. It was validated with a sample of 142 teachers. It was demonstrated that there is a causality relationship of the variables, given by the coefficient of determination R2, the level of influence of the independent variables: levels of reflection and advantages; on the dependent variable, communicative competencies; the level of influence of the exogenous variable: advantages with the endogenous variable: levels of reflection in teachers. A high degree of construct validity was shown through the fit indices of the confirmatory factorial analysis.

Keywords:

Teaching; Virtuality; Podcast; methods; Skills and communication

Introduction

The Covid-19 pandemic has challenged the teaching competences of educators, who have been forced to replace face-to-face instruction with online instruction (Polyakova & Galstyan-Sargsyan, 2021). In this context, a new way of assuming academic and work relationships is generated, which has led to the emergence and diffusion of different digital tools, especially those designed to transfer knowledge in an agile, fun, and motivating way (Craig et al., 2021; Peled, 2021; Samuel-Azran et al., 2019). In this sense, various strategies have been promoted in the development of virtual learning sessions (Polyakova & Galstyan-Sargsyan, 2021), which for many teachers have caused methodological inconsistencies due to the lack of preparation of the educator in the use of digital content (Almendingen et al., 2021; Saeedakhtar et al., 2021), and the lack of connectivity has led to the use of podcasts as an educational tool to generate inclusive and informal learning (Ifedayo et al., 2021).

The use of podcasts in secondary education has been an alternative mainly used to transfer knowledge and develop some dimensions of communicative competence (Basenko & Baskakova, 2021; Borja-torresano et al., 2020). This tool has the advantage of being able to be used online and repeated several times according to the user's time availability (Nalendra et al., 2020; Strickland et al., 2021). In addition, this technological resource allows teaching in a direct and secure way (Strickland et al., 2021), is accessible, free, massive, allows education (Basenko & Baskakova, 2021), and supports students with or without special needs (Alves et al., 2018). In conclusion, it has scientific and social knowledge, can be used in schools, and promotes critical and reflective attitudes (Celaya et al., 2020).

Review Literature

Regarding the literature review, studies have been found that focus on podcasts in the English language (Nalendra et al., 2020; Saeedakhtar et al., 2021), highlighting their effectiveness in oral skills, vocabulary, fluency, and reading comprehension, as well as promoting autonomous learning based on good teaching practices (Almendingen et al., 2021; Barrios-Rubio, 2021; Carson et al., 2021; Drew, 2017; Lowe et al., 2021). Podcasts assume a leading role in message production, inference, and reflection, as well as promoting collaborative work that stimulates critical thinking and allows for the exchange of ideas (González Conde et al., 2021; Saeedakhtar et al., 2021a), proving to be motivating and knowledge-generating.

There are gaps in the use of podcasts in the educational field in relation to knowledge transfer and good practices developed at the institutional level (Drew, 2017). Additionally, there is a need to explore the pedagogical and cost benefits compared to other methods (Almendingen et al., 2021). It is necessary to demonstrate that this learning tool is effective in the teaching-learning process (McNamara & Drew, 2019), and finally, there is a deficit of studies oriented towards the use of podcasts in the early, primary, and secondary educational stages in scientific production (Celaya et al., 2020). As can be seen, we have not found any studies on an instrument that measures the perception of podcasts from their advantages, communicative competencies, and levels of reflection in their perceived representation by teachers in digital pedagogical assignments used in distance education, which will contribute to its significant impact on practices and results obtained in distance education.

Therefore, this research responds to an initiative generated by the Ministry of Education of Peru in its proposal "Aprendo en casa" and "Experiencias de aprendizaje" carried out in 2020 and 2021 (R.M. No 160-2020-MINEDU, 2020), which has not been evaluated in terms of its impact and use at the level of teachers. Thus, this study examined how in-service secondary education teachers perceive the different formats of podcast assignments in the context of virtual learning during the Covid-19 pandemic. This study aims to validate an instrument that can measure the advantages, levels of reflection, and communicative competencies of this tool and explore the success of professional development efforts for blended or hybrid teaching.

Perception of the Podcast from the Educational Context

The podcast allows for a temporal organization of sounds, through voice audio files that are uploaded to different platforms with elements, rhythmic sequences, and meaning that can be stored and shared (Özperçin & Günay, 2020). It is also considered a digital tool created for native speakers that is available for free on the internet for asynchronous use (Basenko & Baskakova, 2021). Unlike podcasting, which tends to be more complex and involves technology, its dissemination is massive without considering a specific recipient or defined purpose (Pareja Aparicio et al., 2019). On the other hand, educational podcasts allow for the transmission of authentic, relevant, and multifunctional content, with defined purposes (Özperçin & Günay, 2020). As a result of the pandemic, teachers have explored different ways to supplement face-to-face teaching, which has allowed the use of this digital tool in the development of learning experiences carried out during the lockdown and to be aware of the benefits of this resource (González Conde et al., 2021).

The Podcast and the Levels of Reflection in the Development of Communicative Competences.

The use of educational multimodality incorporates diverse ways of seeing communication, as well as semiotic resources and organizational channels of the practice of meaning-making with varied methods, resources, and learning discourses, as a complementary domain that supports teaching in schools, fostering autonomy in students (Özperçin & Günay, 2020). Therefore, the use of podcasts promotes active learning, to develop communicative competence from knowing how to listen through stories to develop dialogic, argumentative, and inquiry-based capacity (Carson et al., 2021). As well as learning vocabulary, phonetics, and grammar (Basenko & Baskakova, 2021) and other auditory comprehension skills and motivation (Mirza Suzani, 2021; Saeedakhtar et al., 2021) and increasing vocabulary for better oral, fluent, and precise expression (Özperçin & Günay, 2020; Taylor & Blevins, 2020; Yeh et al., 2021). Consequently, communicative skills such as listening, speaking, writing, and reading (García-Herrera & Erazo-Álvarez, 2020) can be enhanced through the use of podcasts, generating greater autonomy in the development of communication skills (Özperçin & Günay, 2020).

The Advantages of the Podcast in the Levels of Reflection.

Reflection allows for internalizing and interpreting learning experiences, enabling the transfer of knowledge and experiences (Lowe et al., 2021), fostering a dialogic space to link personal experience with the conceptual understanding of pedagogical aspects used by teachers (Carson et al., 2021). Podcasts are useful for providing student-centered instruction, allowing for flexibility in learning at their own pace and convenience, with their descriptive study revealing the effect of teacher behavioral value as a partial mediator of podcast acceptance (Ifedayo et al., 2021).

A study highlighted that informal instruction strategies should possess positive psychological interventions (Dreer, 2021) that allow for these levels of internalization, reflection, and discussion. Furthermore, interprofessional collaboration is considered to foster good educational practices for their use and benefit (Almendingen et al., 2021).

The Advantages of Podcasting in the Development of Communicative Competences

Language is a tool that allows us to interact with other people, understand and construct reality; this is achieved with communicative competences developed in school, social, and cultural life, which complement and feedback each other (MINEDU, 2016), to achieve clear and correct communication in different contexts (Segovia et al., 2013; Martinet et al., 2004). In this sense, the use of podcasts proves to be effective and favorable, as it allows students to listen asynchronously, thereby improving vocabulary mastery (González Conde et al., 2021; Saeedakhtar et al., 2021). Another study highlighted those podcasts make time and space more flexible and agile, energizing educational content, in addition to using other resources (González Conde et al., 2021). Likewise, this educational tool is used interdisciplinary and favors the exercise of auditory comprehension, oral and written expression, promoting student participation and enthusiasm (González Conde et al., 2021; Green et al., 2020; Mirza Suzani, 2021; Romero-Velásquez et al., 2020).

In this sense, variables have been identified around the results of various investigations in the educational context, as presented in Table 1.

Table 1:
Studies on different variables

Variable	Researches
The Podcast and the levels of reflection	Authors (Özperçin & Günay, 2020), (Carson et al., 2021), (Basenko & Baskakova, 2021), (Mirza Suzani, 2021), (Saeedakhtar et al., 2021), (Taylor & Blevins, 2020), (Yeh et al., 2021), (García-Herrera & Erazo-Álvarez, 2020).
Advantages of the Podcast	Authors (Lowe et al., 2021), (Carson et al., 2021), (Ifedayo et al., 2021), (Dreer, 2021), (Almendingen et al., 2021)
Podcast communicative competences	(MINEDU, 2016), (Segovia et al., 2013; Martinet et al., 2004), (González Conde et al., 2021; Saeedakhtar et al., 2021), (Green et al., 2020), (Mirza Suzani, 2021), (Romero- Velásquez et al., 2020).

Source: Own elaboration

Methodology

This work adopted a non-experimental study, with a survey design, to validate a psychometric instrument of the Podcast, based on the experience carried out by high school teachers in virtual education. They completed an online questionnaire between December 9th to 16th, 2021, generated by the school closure due to confinement. To proceed accurately with the statistical analysis, 150 participants completed the survey, and some were excluded because they had at least one missing response in the variables of interest. Consequently, the final sample was formed by 142 teachers. Participation was voluntary, and respondents were recruited through an email invitation, WhatsApp, and advertising on other social platforms. Regarding inclusion criteria, it was decided not to filter by discipline since the podcast has been used for educational purposes interdisciplinarity in all curricular areas, regardless of experience with podcasts or not.

The unintentional snowball sampling was composed of 142 teachers from Regular Basic Education of secondary level, between the ages of 20 to 65 years, mainly from Communication (36.6%); Mathematics (16.9%), Social Sciences (13.0%) and Science and Technology (7.8%) areas, belonging to the public sector (79.0%) and the private sector (20.9%), whose employment status is permanent (56.2%) and contractual (43.7%). Regarding years of service, mainly

in the studied sample, it corresponds to teachers with 26 to 30 years of service (30.7%); between 11 to 15 (16.9%); between 21 to 25, and 0 to 5 years of service (15.0%); from 6 to 10 years (12.4%), and between 16 to 20 years (9.8%).

The instrument was applied in December 2021.

For this reason, demographic data have been identified, presented in Table 2.

Table 2:
Demographic data

Sex	Age	Frequency	Percentage
Female	20 to 24 years old	4	2.8%
	25 to 29 years old	3	2.1%
	30 to 35 years old	9	6.3%
	36 to 40 years old	10	7.0%
	41 to 45 years old	14	9.9%
	46 to 50 years old	17	12.0%
	51 to 55 years old	22	15.5%
	56 to 60 years old	9	6.3%
	61 to 65 years old	2	1.4%
	Sub Total	90	63.4%
Male	20 to 24 years old	2	1.4%
	25 to 29 years old	1	0.7%
	30 to 35 years old	8	5.6%
	36 to 40 years old	5	3.5%
	41 to 45 years old	2	1.4%
	46 to 50 years old	8	5.6%
	51 to 55 years old	13	9.2%
	56 to 60 years old	10	7.0%
	61 to 65 years old	3	2.1%
	Sub Total	52	36.6%
	Total	142	100.0%

Source: Own elaboration

Hypothesis

The Bootstrapping technique was employed for hypothesis testing, which involved extracting a large number of bootstrap resamples (5,000) with replacement from the original sample and then estimating the model parameters for each bootstrap resample. A significance level of $p < 0.05$ was set, indicating that any result with a p-value less than 0.05 is considered statistically significant.

The tested hypotheses were as follows:

H1: The Podcast and the reflection levels -> Communication competences podcast

H2: Podcast Advantages -> The Podcast and the reflection levels

H3: Podcast Advantages -> Communication competences podcast

For each hypothesis, the results of sample means, standard deviations, statistical t-values, and corresponding p-values were presented. The p-values below 0.05 indicated statistically significant relationships between the variables.

The Instrument

They reviewed the literature of various authors with the purpose of identifying the variables to be studied. Subsequently, several items were developed, ensuring they were in line with the reviewed and consulted literature. Finally, they underwent the statistical procedure, resulting in the establishment of the instrument (questionnaire). The following presents three variables:

Variable: Advantages of Podcasting (PCV)

This online tool is flexible in terms of time and can be replayed based on the user's availability. In addition to being secure and direct, it's accessible, free, and valuable for educating students, including those with special needs. In summary, this tool contributes scientific and social knowledge; it's suitable for educational settings (Nalendra et al., 2020; Strickland et al., 2021; Basenko & Baskakova, 2021; Alves et al., 2018; Celaya et al., 2020).

Variable: Podcasting and the Development of Communication Competencies (PCDCC)

Language facilitates interaction and understanding. The communication skills learned in education and society enable effective communication. Podcasts enhance vocabulary and communication, offering flexibility in learning. They also invigorate education and promote auditory, oral, and written skills, exciting students (MINEDU, 2016; González Conde et al., 2021; Saeedakhtar et al., 2021; Green et al., 2020; Mirza Suzani, 2021; Romero-Velásquez et al., 2020).

Variable: Podcasting and Levels of Reflection (PDNR)

Podcasts allow for reflection on learning experiences, enabling the transfer of knowledge and experiences and proving useful for flexible and personalized learning. Additionally, they foster interprofessional collaboration, which enhances educational practices, promoting critical and reflective attitudes (Lowe et al., 2021; Carson et al., 2021; Ifedayo et al., 2021; Almendingen et al., 2021; Celaya et al., 2020).

TITLE: MEANINGS AND PERSPECTIVES OF TEACHERS ON THE USE OF PODCASTS IN VIRTUAL EDUCATION
PROPOSED TITLE: TEACHERS' PERCEPTIONS OF PODCAST USAGE IN VIRTUAL EDUCATION GENERAL OBJECTIVE: To examine teachers' perceptions of podcast usage in the context of virtual education during the Covid-19 pandemic.

VARIABLE	DIMENSION	ITEMS
PODCAST AS AN EDUCATIONAL RESOURCE	ADVANTAGES Authors: Strickland, 2020; Basenko, 2020; Evtyugina y Volkova, 2020; Borja et al., 2020.	1.-Enables asynchronous listening to educational content.
		2.-Allows for an engaging and dynamic listening experience with educational content.
		3.-Has a multifunctional nature.
		4.-Facilitates rapid and widespread dissemination.
		5.-Enables the exploration of current and diverse topics.
		6.-Enhances the understanding of the content.
		7.-Improves students' performance.
		8.-Identifies students' needs and interests.
		9.-Creates a conducive learning environment.
	BENEFITS Authors: Besser, 2021; Strickland, 2020, Basenko, 2020; Romero, 2020.	1.-Enables the achievement of educational objectives for both teachers and students.
		2.-Provides direct and secure teaching to students.
		3.-Improves socio-cultural and communicative skills.
		4.-Develops interdisciplinary skills, approaches, and competencies.
		5.-Promotes autonomy in learning.
		6.-Contributes to the formation of more committed students.
		7.-Deepens students' understanding of the learning material.
		8.-Reduces students' anxiety.
		9.-Benefits students in developing project management skills.
	MEDIATED BY TECHNOLOGY Authors:Strickland, 2020; Basenko, 2020; Romero, 2020; Hitchcock et al., 2021; Makina, 2020.	1.-It is an easy-to-use tool.
		2.-It is ubiquitous, agile, and interactive.
		3.-Leverages the use of ICTs in the teaching-learning process.
		4.-Stimulates dialogue, reflection, and analysis of the social context.
		5.-Allows for an expansion of knowledge in the use of technology.
		6.-It is a stimulating and effective tool.
		7.-Introduces new technological skills.
		8.-Enables content reuse and sharing to strengthen instruction.
		9.-Contributes to the construction of social networks.
		10.-Reduces the technological gap.
	IN THE DEVELOPMENT OF TEACHING STRATEGIES. Authors:Taylor, 2020; Blevins, 2020; Basenko, 2020; Peled, 2021; Andrade, 2020; Moreto 2020; Borja et al., 2020; Mascaro y Ulli, 2020; Franklin, 2007; Unisa, 2015.	1.-Facilitates interactions with teachers and peers on substantive matters.
		2.-Encourages frequent, timely, and constructive feedback.
		3.-Allows for scaffolding. Strengthens the teaching of complex and difficult topics for students.
		4.-Provides opportunities for the development of critical thinking.
		5.-Enhances students' cognitive skills.
		6.-It is a highly visual and auditory learning method.
		7.-Diversifies the teacher's didactic strategies.
		8.-Potentially allows for multitasking.
		9.-Enables innovation in pedagogical practices.
		10.-Consolidates students' knowledge.
	MEASURED BY ONLINE LEARNING Authors:Adawiyah et al., 2021, Richardson 2006	1.-I feel facilitated by online learning.
		2.-I can develop myself, be creative, and innovative with online learning.
		3.-With online learning, I become technologically literate.
		4.-I often have problems using technology.
		5.-I am familiar with the podcast tool.
		6.-It allows for secure content delivery to students.
		7.-Motivates active learning in students through mobile technology.
		8.-Contributes to meeting students' learning expectations as a supportive technology.
	PODCAST ASSIGNMENT WITH LEVELS OF REFLECTION. Authors:Hitchcock et al., 2021; Evtyugina y Volkova, 2020.	1.-Enables personal reflection.
		2.-Allows for information synthesis.
		3.-Connects theory with reality.
		4.-Enables the generation of consistent and relevant arguments.
		5.-Reflects on the meaning of the podcast.
		6.-Reflects on the construction of the podcast.
		7.-Allows for error management.
		8.-Stimulates argumentative discussions among students.
		9.-Enables feedback on learning.
		10.-Considers ethical norms during communication.
	THE PODCAST AND THE DEVELOPMENT OF COMMUNICATION COMPETENCES Authors:Taylor, 2020; Blevins, 2020; Basenko, 2020; Peled, 2021; Andrade, 2020; Moreto 2020; Borja et al., 2020	1.-Improves listening and speaking skills.
		2.-Develops reading and writing skills.
		3.-Fosters vocabulary expansion.
		4.-Enhances pronunciation and fluency in students' oral expression.
		5.-Stimulates oral participation of students.
		6.-Improves intonation and rhythm in students' oral expression.
		7.-Allows for planning the writing process in a text.
		8.-Enables the observation of spelling and grammatical errors in content scripts.
		9.-Facilitates the use of synthesis and organization skills in content script writing.
		10.-Stimulates imagination in students' content script compositions.

The applied questionnaire was developed by the research team. The study variables were evaluated using a five-point Likert scale indicating the degree of agreement or disagreement with the following statements, where (1) is Totally disagree (2) Disagree, (3) Neither agree nor disagree (4) Agree, and (5) Totally agree.

The estimated reliability reached values higher than 0.700 which can be considered acceptable for the Omega and Cronbach's Alpha coefficient. (see Table 3)

Table 3:
Frequent scale reliability statistics

	McDonald's ω	Cronbach's α
Estimate per point	0.978	0.978
95% CI lower limit	0.973	0.972
95% CI upper limit	0.983	0.982

Source: Own elaboration

Analysis

For the exploratory analysis, the variables were processed using a Likert scale, as evidenced by the consistency and magnitude of each item in the instrument. Therefore, these items are correlated, resulting in a Cronbach's alpha of 0.978. The McDonald's omega coefficient is an appropriate measure of reliability, and if the principle of equivalence is not applied, a confidence interval higher than 0.983 will be obtained as a result.

Results

First, in the validation and standardization of the applied instrument, a preliminary test was conducted with 50 study units before the definitive application of the instrument, with the purpose of establishing an adequate discrimination of the instrument items and their corresponding factor, according to the theoretical construct used. Then, a preliminary Exploratory Factor Analysis (EFA) was performed to establish an adequate composition of the items with their corresponding factors, using JASP statistical software v.0.16.

The result of the Kaiser-Meyer-Olkin (KMO) test was 0.912, which is acceptable. The EFA test was satisfactory, using the following criteria: orthogonal rotation (varimax), number of factors: manual, ordering of factor loads by variable criterion and with the output option for each item greater than 0.500 (see Table 4).

Table 4:
Common factor items – factor loadings

	Factor 1	Factor 2	Factor 3	Unicity
PCV1	0.706			0.391
PCV2	0.710			0.391
PCV3	0.764			0.396
PCV4	0.642			0.508
PCV5	0.735			0.350
PCV6	0.723			0.405
PCV7	0.630			0.497
PCV9	0.641			0.447
PDNR3			0.641	0.466
PDNR5			0.613	0.489
PDNR6			0.648	0.559
PDNR7			0.743	0.373
PDNR8			0.686	0.447
PCDCC1		0.742		0.339
PCDCC2		0.601		0.389
PCDCC3		0.551		0.533
PCDCC4		0.668		0.359
PCDCC5		0.740		0.266
PCDCC6		0.734		0.295
PCDCC7		0.813		0.247
PCDCC11		0.620		0.426

Note: The rotation method applied is varimax.

Second, regarding the results of the model, Confirmatory Factor Analysis (CFA) tests were conducted as an analytical condition for the factors, and its application contributed to determining the consistency of the proposed model.

The Comparative Fit Index (CFI) had a value of 0.904, which resulted in a suitable fit. On the other hand, the Tucker-Lewis Index (TLI) is considered suitable for the model if it is ≥ .90 according to Keith (2015, p. 312). However, the obtained result was 0.891, which is below the suggested threshold by Tucker-Lewis. (see Table 5)

Table 5:
Fit Index

Index	Value
Comparative Fit Index (CFI)	0.904
Tucker-Lewis Index (TLI)	0.891

Source: Own elaboration

Third, different authors suggest various criteria for validating fit measures, such as the case of RMSEA (Root Mean Square Error of Approximation), where values equal to or less than 0.05 would be valid (Brown, 2015, p. 72); however, it should be noted that authors such as Jöreskog and Sörbom (1996) suggest values of $P > .50$, while the obtained result is 0.082.

Regarding the SRMR measure, a value of ≤ .08 for good fit is intuitively attractive, although ≤ .06 can be a better

criterion (Keith, 2015, p. 312), and the obtained result is 0.071, which is acceptable according to the criteria explained above. Table 6 presents the achieved fit measures.

Table 6:
Fit Index

Metrics	Value
Root mean square error of approximation (RMSEA)	0.082
RMSEA 90% IC lower limit	0.070
RMSEA 90% IC upper limit	0.095
Standardized Root Mean Square Residual (SRMR)	0.071

Source: Own elaboration

Fourth, the structural design of the proposed model is suitable for Structural Equation Modeling (SEM) based on variances (PLS), using SmartPLS (v.3.3.3) for analysis (Ringle et al., 2015). The external loading test represents the contribution of the indicators to the idea of the theoretical construct so that the measurement model is composed of the indicators and their paths and corresponding factors. The values obtained are called external loadings, and they could range from 0 to 1, so the closer they are to unity, the stronger they are to explain the validity of the proposed model. Table 7 presents the valid factors for the model, which will be represented as independent variables:

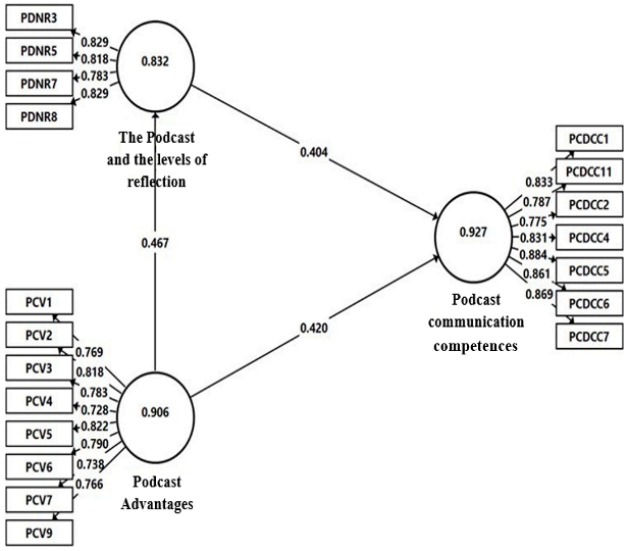
Table 7.
External loadings PLS SEM

	The Podcast and the levels of reflection	Communication competences podcast	Advantages of Podcast
PCDCC1		0.833	
PCDCC11		0.787	
PCDCC2		0.775	
PCDCC4		0.831	
PCDCC5		0.884	
PCDCC6		0.861	
PCDCC7		0.869	
PCV1			0.769
PCV2			0.818
PCV3			0.783
PCV4			0.728
PCV5			0.822
PCV6			0.790
PCV7			0.738
PCV9			0.766
PDNR3	0.829		
PDNR5	0.818		
PDNR7	0.783		
PDNR8	0.829		

Note:
PCDCC: Podcasting and The Development of Communication Competencies
PCV: Advantages of Podcasting
PDNR: Podcasting And Levels Of Reflection
Source: Own elaboration

The consistency of the model was expressed through Cronbach's alpha to establish the degrees of reliability of the indicators, the results exceed 0.700, being acceptable values for the structural model (Hair et al., 2010). (see Figure 1)

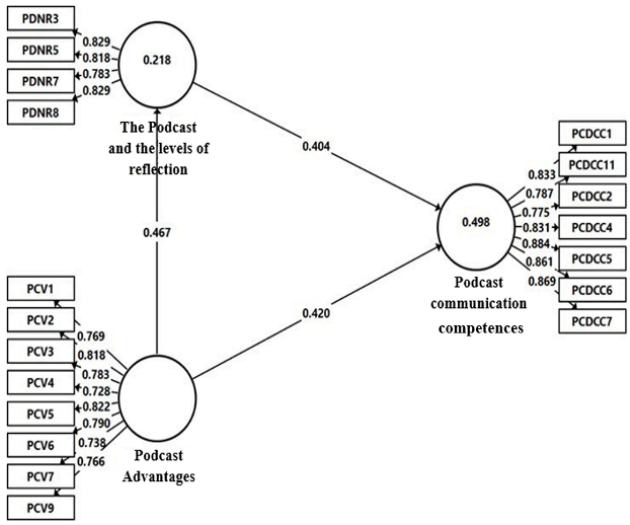
Figure 1.
Cronbach alpha of the SmartPLS model



In the measurement model, reliability and validity are obtained by calculating the Partial Least Squares (PLS) algorithm, for which different path coefficients are obtained. The results are explained based on the R2, which reflects the goodness of fit to establish the relationship between the proposed factors in the model. In this sense, communicative competencies through the use of the Podcast are positively influenced by the level of reflection on its use and the advantages that come with its proper use, with a percentage of 49.8%, and the most influential variable being the advantages of its use.

The model also allows for the level of influence of the exogenous variable: advantages in the use of the Podcast on the endogenous variable: levels of reflection in teachers, based on the R2 coefficient, which is 0.218, indicating that its influence would be 21.8%. (see Figure 2).

Figure 2.
R2 from SmartPLS model



For the reliability and validity of the construct, correlation coefficients are taken into account. The values for the Cronbach's alpha, according to the proposed model, are between 0.832 to 0.927. Thus, they are statistically significant since they meet the minimum threshold recommended by F. Hair Jr et al. (2014). Regarding the Average Variance Extracted (AVE) values, they range from 0.604 to 0.698, which corresponds to the value suggested by Chong (2013) of 0.650. For Composite Reliability, the application of Bagozzi & Yi (1988) and Hair et al. (2012) is suggested. According to these authors, if the results are greater than 0.6, then high levels of internal consistency reliability are demonstrated for each of the variables. In this sense, the results obtained are between 0.888 to 0.942. The (rho_A) coefficient is used to verify the reliability of the values obtained in the construction and design of the PLS. The recommended results should be greater than 0.7. According to the data shown in Table 8, all coefficients exceed 0.700, demonstrating a high level of reliability.

Table 8.

Reliability and construct validity

	Cronbach alpha	Rho_A	Composite reliability	Average variance extracted (AVE)
The Podcast and the reflection levels	0.832	0.835	0.888	0.664
Communication competences podcast	0.927	0.929	0.942	0.698
Podcast Advantages	0.906	0.910	0.924	0.604

Source: Own elaboration

The discriminant or divergent validity , shows that constructs should not have any relationship. To establish this criterion of analysis, two methods were used. The first is the Fornell and Larcker method (Fornell & Larcker, 1981), which suggests that the square root of AVE in each variable, its resulting values should be greater than the results of the correlation between the studied variables. The results indicate that there is discriminant validity, therefore, it meets this requirement (see Table 9).

Table 9.

Fornell-Larcker criterion analysis for discriminant validity and verification

	The Podcast and the reflection levels	Communication competences podcast	Podcast Advantages
The Podcast and the reflection levels	0.815		
Communication competences podcast	0.600	0.835	
Podcast Advantages	0.467	0.609	0.777

Source: Own elaboration

Another method for verifying discriminant validity is the one proposed by Henseler, C. M. Ringle, and M. Sarstedt (Henseler et al., 2015), called Heterotrait-Monotrait (HTMT). The results obtained are justified because their values are below the conservative threshold of 0.850 proposed by Franke and Sarstedt (2019), although they can go up to 0.90 if the constructs are conceptually similar (Benitez et al., 2020; Henseler et al., 2015; Ogbeibu et al., 2018). According to the results, the values obtained are below 0.85 (see Table 10).

Table 10.

Heterotrait-Monotrait Criterion - HTMT for verification and discriminant validity

	The Podcast and the reflection levels	Communication competences podcast	Podcast Advantages
The Podcast and the reflection levels			
Communication competences podcast	0.681		
Podcast Advantages	0.524		0.658

Source: Own elaboration

The hypothesis testing was conducted using the Bootstrapping technique, which is a process of extracting a large number of bootstrap resamples (5,000) with replacement from the original sample and then estimating the model parameters for each bootstrap resample. Considering the significance level for the P Value ($p < 0.05$), all the hypotheses proposed are accepted. (see Table 11).

Discussion

The coefficient of determination indicates the predictive accuracy of the model and the combined effect of the independent variables on the dependent variable. Overall, the SEM analysis results have supported a significant influence among all proposed hypotheses. From a practical perspective, this study has shown that H1 The Podcast and reflection levels positively influence the development of communication competencies, finding that they enhance pronunciation, fluency, intonation, and rhythm in oral expression; foster writing skills and increase vocabulary with a high factorial loading. Similarly, it was found that the podcast is a didactic resource capable of promoting oral, reading, and writing skills; in addition to being a useful digital tool for feedback, correction, and enhancement of communication skills (Loja-Gutama et al., 2020; Taylor & Blevins, 2020; Mirza Suzani, 2021; Yeh et al., 2021). It also creates need for discipline and commitment in its listeners (Özperçin & Günay, 2020; Shamburg, 2020), likewise leading to self-reflection and monitoring of their performance regarding technical challenges, language difficulties, and progression. Moreover, it

Table 11.

Hypothesis testing - Bootstrapping

	Original sample (O)	Sample mean (M)	Standard deviation (STDEV)	Statistical t (|O/STDEV|)	P Value
H1 The Podcast and the reflection levels -> Communication competences podcast	0.404	0.415	0.087	4.637	0.000
H2 Podcast Advantages -> The Podcast and the reflection levels	0.467	0.476	0.083	5.663	0.000
H3 Podcast Advantages -> Communication competences podcast	0.420	0.408	0.121	3.471	0.000

Source: Own elaboration

contributes to developing their oral communication skills in a thoughtful manner (Özperçin & Günay, 2020) by allowing them to record their voices, listen to themselves, and edit the final version of their multimodal communication.

Regarding H2 the advantages of Podcast positively influence the levels of reflection perceived by teachers, their articulation with theory and reality is identified with high values, followed by feedback and metacognition. Thus, it is affirmed that the use of podcasts favored the creation of spaces for discussion, reflection, and collaboration in group tasks that were more attractive because they started from reality (Almendingen et al., 2021). In addition to this, the result of using the podcast as a social, technological, and discursive space where narratives and stories are shared, which were analyzed collaboratively and critically, moved to an emotional and conceptual level (Carson et al., 2021), must be prioritized for better teaching. Therefore, in continuous professional development, recorded dialogues and reflective responses should be used to examine transformations in thinking and action of their teaching beliefs and practices (Lowe et al., 2021), which benefit their successful implementation of the podcast in an enjoyable and flexible way (Besser et al., 2021; González Conde et al., 2021; Ifedayo et al., 2021).

H3 The advantages of the Podcast have a positive influence on the development of communication competencies, which are reflected in the factor loadings: PCV5, the Podcast allows for the treatment of current and diverse topics; PCV2, the Podcast allows for enjoyable and dynamic listening of educational content, and PCV6, the Podcast deepens the understanding of the content, which are the perceptions assumed by teachers. These findings are similar to other studies, as it flexibilizes space and time, is an innovative, motivating, and effective resource that enhances linguistic, social, and personal formation (González Conde et al., 2021; Mirza Suzani, 2021). Additionally, it strengthens auditory skills of various texts with high educational and satisfaction value (Green et al., 2020; Mirza Suzani, 2021; Romero-Velásquez et al., 2020).

Furthermore, if internalized collaboratively, it improves auditory comprehension achieving a vocabulary domain, because audios can be listened to and adapted asynchronously. (Saeedakhtar et al., 2021).

The document contributes by delivering an instrument that measures the perceived value of podcast perception and value as a means of reflecting on the methodology employed and its effect on student learning from the pandemic context. Coinciding with the fact that educational podcasts need to seek theoretical frameworks that allow the development of these in unique learning environments (McNamara & Drew, 2019). A second contribution is reflected in the adoption of the podcast tool by teachers in the Southern region of Peru, showing that the three variables have significant representativeness (Basenko & Baskakova, 2021; Borja-torresano et al., 2020).

The results of this research confirm outstanding factors to determine the usefulness and value of teachers in the significant use of podcasts. Therefore, by implementing podcasts as an online teaching and learning strategy, it leads to teachers being informed about the characteristics of the tool, its technical problems, as well as its usefulness, so that they feel secure and can plan its use, coinciding in the sense that educational podcasts are feasible if they have a positive impact on learning (McNamara & Drew, 2019).

The practical implications of this study were presented in the interaction among the three variables, as it represents the first investigation that examined the advantages of podcast use, the value of communicative competence, and levels of online reflection from the perception of secondary education teachers. In general, we believe that these results can be particularly interesting to shed more light on the validity of the instrument model. However, the benefit of this task offered a variety of practical transferable skills related to technology and

focused on teaching-learning and the need to quickly adapt and respond immediately to the difficulties of digital gaps mediated by technology. This digital tool demands double learning, first to familiarize oneself with the use of the device, second to make didactic use of the podcast that allows for meaningful learning (Ifedayo et al., 2021).

This study provided a limited sample of the perceptions of public and private sector teachers, mostly from the Communication area. Therefore, it is suggested that it can be replicated with other groups of teachers and analyze the effects of the perceptions and subjective value of teachers on their own experiences in the use of the podcast in the educational context. Although a significant influence was found among the three factors PCDD, PDNR, and PCV, the study did not measure different theoretical frameworks used to support the methodology employed in the use of podcasts in virtual learning environments. Therefore, it would be advisable to analyze these cognitive, social, motivational, and other experiences (McNamara & Drew, 2019). As well as, to evaluate the impact of hybrid or blended teaching on the use of podcasts at both the student and teacher levels.

Conclusion

The results of this study confirm the validity and reliability of the psychometric instrument for measuring the advantages of using podcasts, levels of reflection, and communicative competences, based on the methodological and practical aspects exercised by high school teachers in virtual education. The findings confirm that the communicative competences factor through the use of podcasts is positively influenced by the level of reflection and advantages. A causality relationship is also observed between the "advantages" variable and the "levels of reflection" variable, referred to the teachers.

This questionnaire has multiple uses, among them, it can be highlighted that it is useful for research that seeks to evaluate the impact of using podcasts, identify gender-related factors that influence their use and advantages.

Until now, there is no similar quantitatively and statistically validated instrument in the literature. This construct analysis facilitated the exploration and clarification of the podcast as a didactic resource in the educational field and by demonstrating versatility, it opens up the need to focus on new studies on the theoretical assumptions of the educational podcast. It is suggested that future research could evaluate the degree to which podcasts contribute to learning, according to the standards and competencies developed in educational curricula, as well as the methodologies employed by teachers and their effectiveness and impact.

References

Agaton, C. B., & Cueto, L. J. (2021). Learning at home: Parents' lived experiences on distance learning during COVID-19 pandemic in the Philippines. *International Journal of Evaluation and Research in Education (IJERE), 10*(3), 901. https://doi.org/10.11591/ijere.v10i3.21136

Almendingen, K., Torbjørnsen, A., Sparboe-Nilsen, B., Kvarme, L. G., & Saltyte Benth, J. (2021). Small Group Student-Produced Podcasts Were Favoured as Assignment Tool for Large-Scale Interprofessional Learning: An Exploratory Study Among Health, Social Care, and Teacher Education Program. Frontiers in Education, 6. https://doi.org/10.3389/feduc.2021.622716

Alves, K. D., Kennedy, M. J., Kellems, R. O., Wexler, J., Rodgers, W. J., Romig, J. E., & Peeples, K. N. (2018). Improving Preservice Teacher Vocabulary Instruction: A Randomized Controlled Trial. *Teacher Education and Special Education: The Journal of the Teacher Education Division of the Council for Exceptional Children, 41*(4), 340–356. https://doi.org/10.1177/0888406417727044

Bagozzi, R. P., & Yi, Y. (1988). On the evaluation of structural equation models. *Journal of the academy of marketing science, 16*(1), 74-94.

Barrios-Rubio, A. (2021). Radio, music and podcast in the consumption agenda of colombian adolescents and youth in the digital sonosphere. *Communication and Society, 34*(3), 31–46. https://doi.org/10.15581/003.34.3.31-46

Basenko, G., & Baskakova, V. (2021). Podcasts in the teaching media space. E3S Web of Conferences, 273. https://doi.org/10.1051/e3sconf/202127312122

Benitez, J., Henseler, J., Castillo, A., & Schuberth, F. (2020). How to perform and report an impactful analysis using partial least squares: Guidelines for confirmatory and explanatory IS research. *Information & Management, 57*(2), 103168.

Besser, E. D., Blackwell, L. E., & Saenz, M. (2021). Engaging Students Through Educational Podcasting: Three Stories of Implementation. Technology, Knowledge and Learning, 0123 789. https://doi.org/10.1007/s10758-021-09503-8

Borja-torresano, S. C., Ulli-flores, W. E., & Of, B. (2020). *Sonia Cecilia Borja-Torresano. 6*, 172–197.

Brown, T. A. (2015). Confirmatory Factor Analysis for Applied Research, (Second edition ed.). The Guilford Press

Cain, J., Cain, S., & Daigle, B. (2021). Constructivist Podcasting Strategies in the 8th Grade Social Studies Classroom:"StudyCasts" Support Motivation and Learning Outcomes. The Social Studies, 1-12. https://doi.org/10.1080/00377996.2021.1934810

Carson, L., Hontvedt, M., & Lund, A. (2021a). Student teacher podcasting: Agency and change. Learning, Culture and Social Interaction, 29(April), 100514. https://doi.org/10.1016/j.lcsi.2021.100514

Celaya, I., Ramírez-Montoya, M. S., Naval, C., & Arbués, E. (2020). Uses of the podcast for educational purposes. Systematic mapping of the literature in WoS and Scopus (2014-2019). *Revista Latina de Comunicacion Social, 2020*(77), 179–201. https://doi.org/10.4185/RLCS-2020-1454

Chong, A. Y.-L. (2013). A two-staged SEM-neural network approach for understanding and predicting the determinants of m-commerce adoption. *Expert Systems with Applications, 40*(4), 1240-1247. https://doi.org/https://doi.org/10.1016/j.eswa.2012.08.067

Dreer, B. (2021). Fostering Well-Being over the Radio? An Empirical Study Investigating the Effects of an Audio Podcast-Based Intervention Program on Student Teachers' Well- Being. *International Journal of Community Well-Being, 4*(4), 603–623. https://doi.org/10.1007/s42413-020-00105-3

Drew, C. (2017). Edutaining audio: an exploration of education podcast design possibilities. *Educational Media International, 54*(1), 48 62. https://doi.org/10.1080/09523987.2017.1324360

F. Hair Jr, J., Sarstedt, M., Hopkins, L., & G. Kuppelwieser, V. (2014). Partial least squares structural equation modeling (PLS-SEM). *European Business Review, 26*(2), 106- 121. https://doi.org/10.1108/EBR-10-2013-0128

Fornell, C., & Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of marketing research, 18*(1), 39-50.

Franke, G., & Sarstedt, M. (2019). Heuristics versus statistics in discriminant validity testing: a comparison of four procedures. Internet Research.

García-Herrera, D. G., & Erazo-Álvarez, J. C. (2020). http://dx.doi.org/10.35381/e.k.v3i1.1016. Podcast Como Recurso Didáctico Para Desarrollar Habilidades Comunicativas, III, 355–375.

Fornell, C., & Larcker, D. F. (1981). Evaluating structural equation models witunobservable variables and measurement error. *Journal of marketing research, 18*(1), 39-50.

Franke, G., & Sarstedt, M. (2019). Heuristics versus statistics in discriminant validity testing: a comparison of four procedures. Internet Research.

García-Herrera, D. G., & Erazo-Álvarez, J. C. (2020). http://dx.doi.org/10.35381/e.k.v3i1.1016. Podcast Como Recurso Didáctico Para Desarrollar Habilidades Comunicativas, III, 355–375.

González Conde, M. J., Prieto González, H., & Baptista Gil, F. (2021). Didáctica del podcast en el programa PMAR. Una experiencia de aula en la Comunidad de Madrid. RIED. *Revista Iberoamericana de Educación a Distancia, 25*(1), 183–201. https://doi.org/10.5944/ried.25.1.30618

Green, K. B., Stuckey, A., Towson, J. A., Robbins, S. H., & Bucholz, J. L. (2020). Special Education Preservice Teacher Knowledge of Mathematics Methods: The Effects of Content Acquisition Podcasts (CAPs). *Journal of Special Education Technology, 35*(3), 145–154. https://doi.org/10.1177/0162643419854494

Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2010). Multivariate Data Analysis.Prentice Hall.

Hair, J. F., Sarstedt, M., Pieper, T. M., & Ringle, C. M. (2012). The use of partial least squares structural equation modeling in strategic management research: a review of past practices and recommendations for future applications. *Long range planning, 45*(5-6), 320-340.

Henseler, J., Ringle, C. M., & Sarstedt, M. (2015). A new criterion for assessing discriminant validity in variance-based structural equation modeling. *Journal of the academy of marketing science, 43*(1), 115-135.

Ifedayo, A. E., Ziden, A. A., & Ismail, A. B. (2021). Mediating effect of behavioural intention on podcast acceptance. *Education and Information Technologies, 26*(3), 2767–2794. https://doi.org/10.1007/s10639-020-10385-z

Jöreskog, K. G., & Sörbom, D. (1996). LISREL 8: User's Reference Guide. Scientific Software International.

Keith, T. Z. (2015). Multiple Regression and Beyond: An Introduction to Multiple Regression and Structural Equation Modeling (Second edition ed.). Taylor & Francis.

Loja-Gutama, B. D., García-Herrera, D. G., Erazo-Álvarez, C. A., & Erazo-Álvarez, J. C. (2020). Podcast como estrategia didáctica en la enseñanza de la expresión oral y escrita. *Cienciamatria, 6*(3), 167–192. https://doi.org/10.35381/cm.v6i3.395

Lowe, R. J., Turner, M. W., & Schaefer, M. Y. (2021). Dialogic research engagement through podcasting as a step towards action research: A collaborative autoethnography of teachers exploring their knowledge and practice. *Educational Action Research, 29*(3), 429–446. https://doi.org/10.1080/09650792.2021.1908905

Lundström, M., & Lundström, T. P. (2021). Podcast ethnography. *International Journal of Social Research Methodology, 24*(3), 289-299. https://doi.org/10.1080/13645579.2020.1778221

McNamara, S., & Drew, C. (2019). Concept analysis of the theories used to develop educational podcasts. *Educational Media International, 56*(4), 300–312. https://doi.org/10.1080/09523987.2019.1681107

MINEDU (2016). Programa Curricular de Educación Secundaria. http://www.minedu.gob.pe/curriculo/pdf/03062016-programa-nivel-secundaria-ebr.pdf

Mirza Suzani, S. (2021). Investigating the Effect of Podcasting on Iranian Senior Undergraduate TEFL Students' Listening Comprehension Improvement and Motivation. *Asia-Pacific Education Researcher, 30*(5), 395–408. https://doi.org/10.1007/s40299-020-00526-w

Ogbeibu, S., Senadjki, A., & Gaskin, J. (2018). The moderating effect of benevolence on the impact of organisational culture on employee creativity. *Journal of Business Research, 90*, 334-346.

Özperçin, A., & Günay, D. (2020). Utilizing podcasting as a multimodal rehearsal task for fostering communicative competence of pre-service FLE teachers in Istanbul University- Cerrahpaşa. Synergies Turquie, 13, 27–44.

Pareja Aparicio, M., Terol Bolinches, R., & Alonso López, N. (2019). La utilización del podcast como herramienta pedagógica. Caso de estudio de la asignatura Teoría y Ecología de los Medios Audiovisuales. 183–197. https://doi.org/10.4995/inred2019.2019.10366

Resolución Ministerial No 160-2020-MINEDU. Diario Oficial El Peruano (31 de marzo del 2020). https://www.gob.pe/institucion/minedu/normas-legales/466108-160-2020- minedu

Ringle, C. M., Wende, S., & Becker, J.-M. (2015). SmartPLS 3. Boenningstedt: SmartPLS GmbH.

Romero-Velásquez, I. F., García-Herrera, D. G., Erazo-Álvarez, C. A., & Erazo-Álvarez, J. C. (2020). Podcast como recurso didáctico para desarrollar habilidades comunicativas. *Episteme Koinonia, 3*(1), 355. https://doi.org/10.35381/e.k.v3i1.1016

Saeedakhtar, A., Haqju, R., & Rouhi, A. (2021). The impact of collaborative listening to podcasts on high school learners' listening comprehension and vocabulary learning. *System, 101*(May), 102588. https://doi.org/10.1016/j.system.2021.102588

Shamburg, C. (2020). Cases of successful independent educationally oriented podcasters.*E-Learning and Digital Media, 17*(6), 505–520. https://doi.org/10.1177/2042753020946281

Taylor, J. L., & Blevins, M. (2020). COMMcast: Producing podcasts for communication theory. *Communication Teacher, 34*(4), 272–276. https://doi.org/10.1080/17404622.2019.1706756

Yeh, H. C., Chang, W. Y., Chen, H. Y., & Heng, L. (2021). Effects of podcast-making on college students' English speaking skills in higher education. *Educational Technology Research and Development, 69*(5), 2845–2867. https://doi.org/10.1007/s11423-021-10026-3

Appendix 1

Instrument

Advantages of Podcasting (PCV)

PCV1: Podcasting as an advantage allows educational content to be listened to asynchronously.
PCV2: Podcasting as an advantage allows for educational content to be listened to in an enjoyable and dynamic way.
PCV3: Podcasting as an advantage has a multifunctional character.
PCV4: Podcasting as an advantage allows for rapid and massive dissemination.
PCV5: Podcasting as an advantage allows for treatment of current and diverse topics.
PCV6: Podcasting as an advantage deepens understanding of content.

Podcasting and the Development of Communication Competencies (PCDCC)

PCDCC1: Podcasting allows for improvement of listening skills.
PCDCC2: Podcasting allows for improvement of speaking skills.
PCDCC4: Podcasting promotes an increase in vocabulary.
PCDCC5: Podcasting improves pronunciation and fluency in oral expression for students.
PCDCC6: Podcasting stimulates oral participation for students.
PCDCC7: Podcasting improves intonation and rhythm in oral expression for students.
PCDCC11: Podcasting stimulates imagination in script writing for student content.

Podcasting and Levels of Reflection (PDNR)

PDNR3: Podcasting articulates theory with reality.
PDNR5: Podcasting allows for metacognitive reflection.
PDNR7: Podcasting stimulates argumentative discussion among students.
PDNR8: Podcasting allows for feedback on learning.

This page is intentionally left blank.

www.**iejee**.com

iejee

International Electronic Journal of Elementary Education
September 2023, Volume 16, Issue 1, 57-73

Investigating Beliefs of Teachers of Multilingual Learners (MLLs)

Burhan Ozfidan[a,*], Rabia Hos[b]

Received : 13 June 2023
Revised : 21 August 2023
Accepted : 29 September 2023
DOI : 10.26822/iejee.2023.314

[a,*] **Corresponding Author:** Burhan Ozfidan, Florida Gulf Coast University
E-mail: bozfidan@fgcu.edu
ORCID: https://orcid.org/0000-0002-2098-3409

[b] Rabia Hos, University of Rhode Island
E-mail: rabiahos@uri.edu
ORCID: https://orcid.org/0000-0002-3197-9771

Abstract

The purpose of this study is to measure teachers' beliefs and perceptions about multilingual learners (MLLs), MLLs' needs and struggles, and parental involvement of MLL families. We adapted Karabenick and Noda's (2004) survey instrument, which measures the teachers' beliefs, attitudes, practices, and needs related to MLLs. Exploring teachers' beliefs, attitudes, and practices toward MLLs is a first step in understanding the professional needs of teachers of MLLs. The total sample size of the study was 308. In the first phase of the study, Confirmatory Factor Analysis (CFA) was used to test logic and feasibility of the instrument. In the second phase, we used descriptive statistics to explain the basic features of the data of the study and provide a brief summary of the samples and the measures done on this study. The results of this study indicated that although teachers' beliefs about MLLs are slightly negative their overall beliefs are positive. Many teachers while welcoming MLLs in their classrooms, desiring to support them, are unsure of how to do so.

Keywords:

Multilingual Learners, Beliefs of Teachers, Classroom Experience

Introduction

The number of multilingual learners (MLLs) has drastically increased in schools in the United States, now making up more than 10% of the total student population (NCES, 2019). While the majority of MLLs are mostly found in nine U.S. states, many of the smaller states have also had steady increase in the numbers in the past ten years (Batlova & McHugh, 2010). The experiences of MLL in schools are in partly related to the beliefs and perceptions that their teachers hold about them. MLLs will reach their fullest potential and achieve success when teachers have positive beliefs and high expectations for them (McSwain, 2001). Knowledge of second language acquisition processes, self-reported cultural competency of teachers, and teachers' beliefs and perceptions about MLLs in general have an effect on MLLs experience schools (Harrison & Lakin, 2018). In order to "understand teaching from teachers' perspectives we have to understand the beliefs with which they define their work" (Thompson, 1992, p. 129). With the increasing demands of the Common Core State Standards (CCSS), which is a federally mandated initiative

to use common content standards for elementary and secondary students and focusing on both linguistic and academic content knowledge, there is an increased need for all teachers to respond to the needs of the increasing number of MLLs in their classrooms (Valdes, Kibler, & Walqui, 2014; Wright, 2006). Although some studies have focused on teachers' ideologies and beliefs when working with language learners and what their ideologies are rooted in (Kang, 2022; Young, 2014), there has been a paucity in research looking at the various belief systems teachers may hold toward language acquisition, multilingualism, and cultural competency (Karabenick & Noda, 2004). We understand the beliefs of teachers that underpins their professional practice (Martin-Jones, 2009). Examining the teachers' beliefs and perceptions about their own preparedness to meet the needs of MLLs and toward MLLs in general are highly influential in students' success (Karabenick & Noda, 2004; Ozfidan & El-Dakhs, 2023). It is also important to note here that teachers' beliefs are influenced by societal norms and contexts (Walker, Shafer, & Iiams, 2004). If teachers have unexamined negative beliefs toward MLLs, even in most well-meaning teachers may potentially discriminate against MLLs unconsciously. Therefore, the purpose of this study is to understand teachers' beliefs and perceptions about MLLs. In order to do this, we adapted Karabenick and Noda's (2004) survey instrument, which measures "the teachers' beliefs, attitudes, practices, and needs related to MLLs" (p.56). Many teachers while welcoming MLLs in their classrooms, desiring to support them, are unsure of how to do so.

Theoretical Framework

Educators may believe that working with MLLs requires the dedication of too much instructional time and resources, compared with English-monolingual peers (Bailey & Marsden, 2017; Karabenick & Noda, 2004). Teachers often correlate bi- or multi-lingualism with the need for additional support rather than with cognitive, social, cultural, and/or linguistic benefits (Bailey & Marsden, 2017; Butcher, Sinka, & Troman, 2007). This association reinforces a deficit view of bi- or multi-lingualism (Butcher, Sinka, & Troman, 2007). In Bailey and Marsden's (2017) study of seven teachers, only one educator described the ability to speak two languages as an advantage. In considering the drawbacks associated with MLLs, educators commonly identified slow individual academic progress, low student confidence, lack of belonging, and general social disadvantages (Bailey & Marsden, 2017). High-stakes standardized testing and punitive accountability systems that penalize educators can compound teachers' concerns about student achievement and potentially create negative attitudes toward MLLs (Mellom, Straubhaar, Balderas, Ariail, & Portes, 2018).

The researchers adapted Karabenick and Noda's (2004) survey to measure teachers' beliefs regarding their own cultural competence, multilingual learners' needs, and parental involvement. Therefore, theoretical background relevant to the constructs included in Karabenick and Noda's original survey. Karabenick and Noda considered it important to measure teachers' successed goals, owing to their current dominance in the conceptual framework on school motivation. This method differentiates among teachers' focus on mastery versus performance (Midgley, Kaplan, & Middleton, 2001; Midgley, 2002; Nicholls, 1984). Mastery goals highlight involvement in individual improvement and tasks. On the contrary, performance goals highlight students' abilities, which are made more noticeable comparing inter-student. According to Midgley (2002), "a focus on mastery goals is more conducive to the development of intrinsic interest and long-term motivation, whereas a stress on performance goals (especially when they emphasize the avoidance of failure) is linked to poorer performance and negative emotions" (p.142). Multilingual learners (MLLs) are more likely than non- MLLs to suffer from interpersonal comparisons and competition because of language and cultural differences. Thus, MLLs are more likely to thrive in classrooms that are more mastery focused and less performance focused.

Importantly, Karabenick and Noda (2004) determined that increased teacher contact with MLLs was positively correlated to favorable attitudes toward having MLLs in their classrooms. The most positive impressions of MLLs were associated with less experienced teachers in elementary schools (compared with more experienced teachers at the high school-level). The researchers' results also demonstrated that educators with more favorable attitudes toward MLLs generally adopted a mastery approach to instruction versus a performance (or competitive) approach to instruction.

Karabenick and Noda (2004) identified issues of teacher ambivalence in their study, particularly related to the notion of whether or not people of different cultures can work and socialize together, in addition to pinpointing ambivalent educator perspectives about familial support and care. Problematically, 62% of participating teachers believed that the parents of MLLs were less involved in the schools than parents of other students (Karabenick & Noda, 2004). The current study investigated whether teachers with more positive attitudes toward MLLs stressed the more beneficial (i.e., mastery) goals for achievement.

Research on teachers' beliefs

For the purposes of this study, we are using the term belief in its broadest terms, referring to teachers' beliefs as attitudes as part of the cognitive aspects of teaching. This study encompasses multiple factors

of teachers' beliefs about MLLs, including knowledge about language acquisition, attitudes toward language teaching, parental involvement of MLLs, and teachers' own cultural competency. We position ourselves in that teachers' beliefs about teaching and learning are affected by their own experiences and have a great impact on their decisions for teaching and what and how they teach language learners (Abdi & Asadi, 2015; Phillips & Borg, 2009). We begin with an understanding that as teacher educators we need to focus on understanding the beliefs of our teacher candidates so that we can set out to establish programs to shift any negative beliefs that teachers may have toward MLLs and work toward improving the educational experiences of MLLs. We argue that beliefs can shift over time with professional development.

Beliefs are thoughts and judgements that we make about ourselves, others, and the world around us are ideas that are based on our thought processes (Khader, 2012). Beliefs are ideas that are shaped by our experiences (Khader, 2012). Beliefs are defined as the teachers' opinions and their views on teaching and learning (Haney, Lumpe & Czerniak, 1996). Beliefs of teachers not only affect teachers' pedagogy in the classroom, but also their expectations they hold for their students. Studying teacher beliefs is critical in teacher education because beliefs "drive classroom actions and influence the teacher change process" (Richardson, 1996, p. 102). Teachers beliefs have influential implications for MLLs in their classrooms. For example, Rueda and Garcia (1996) found that teachers' beliefs about second language learning and teaching shape their perceptions and judgments, which, in turn, affect students' behavior in the classroom. Additionally, teachers' attitudes toward MLLs affect the classroom interaction between the students and the teacher, which directly affects student success.

Different terms that have been used in the literature about teacher beliefs are beliefs, attitudes, ideologies, perceptions, dispositions, cognition, values, and expectations (Knopp & Smith, 2005). While some researchers focused on multiple definitions of beliefs (Pajares, 1992; Kagan, 1992), others have referred to it as teacher cognition (Borg, 2003; Richards & Lockhart, 1994). Beliefs of teachers can be categorized as implicit and explicit, contextualized in nature, and may change over time (Fives & Buehl, 2012; Mantero & McVicker, 2006). Beliefs can be both subjective and objective and drive teachers' day to day decision making and actions in the classroom (Richards & Lockhart, 1994). Some scholars have argued that teachers' professional knowledge can be considered as belief (Kagan, 1992). Fang (1996) added that beliefs are part of teachers' knowledge. Fenstermacher (1994) proposed that beliefs and knowledge can be used interchangeably.

Teachers' beliefs and perceptions of MLLs

Teachers hold a variety of beliefs about Multilingual Learners (MLLs), bi- or multi-lingualism, and cultural or language diversity. In a study of 729 teachers at a midwestern suburban school district in the United States, Karabenick and Noda (2004) found that 89% of teachers believed that cultural differences enrich the lives of community members (p. 68). Karabenick and Noda (2004) established that teachers more positively oriented toward MLLs in their classrooms may be more likely to believe the following: First language proficiency supports academic success and does not hinder learning another language; bilingualism and bilingual education are beneficial; MLLs should be tested in their first language; lack of fluency in the second language does not indicate a lack of comprehension; and MLLs do not necessarily demand additional time or resources. 75% of surveyed teachers indicated the belief that bilingualism has practical, career-related advantages, and 52% of respondents believed that higher levels of bilingualism develop cognitive skills (Karabenick & Noda, 2004, p. 64).

Bailey and Marsden (2017) interviewed seven educators teaching at predominantly monolingual primary schools in England and learned that teachers held a variety of favorable beliefs about MLLs. For instance, participating teachers expressed that MLLs could contribute to teaching student-peers, that MLLs applied themselves to learning activities, and that MLLs' families possessed a strong cultural work ethic and were interested in their students' well-being and academic achievement. One of the implications of Bailey and Marsden (2017)'s study was incorporating home language into the classroom in order to improve the educational, social, and individual experience of MLLs. These beliefs include that MLLs deserve to fully access the curriculum, that it is appropriate to celebrate diversity by honoring students' cultures and lived experiences, and that MLLs should be welcomed into the classroom environment.

The lens with which teachers view MLLs is often rooted in their beliefs and how this is reflected in their teaching. In one study, preservice teachers working with MLLs characterized the MLLs as "victims" with insurmountable challenges, rather than students with unique skills and strengths (Sugimoto, Carter, & Stoehr, 2017, p. 186). The negative image of MLLs may lead to the "lower expectations" that in-service teachers may have for their language-learning students (Ukpokodu, 2007, p.8).

In another study, Kelly (2018) set out to determine if preservice ESL teachers do, in fact, hold a "deficit view" of English learners (p. 112). Several college students enrolled in an ESL teacher education program drew depictions of their idea of what teachers teaching MLLs should look like, at the beginning and end of

an ESL methods course (Kelly, 2018, p. 110). Most of the preservice teachers depicted direct instruction, in which the ESL students were passive recipients of the lesson (Kelly, 2018, p. 120). Most of the participants did not change their drawings much after completing the course, signalling that their perceptions of MLLs as passive members of the classroom remained the same (Kelly, 2018, p. 124-125).

Teachers' beliefs about multi-lingualism and linguicism

In Young's (2014) study of 46 head teachers in France about their attitudes toward plurilingualism and how their ideologies influence how language policies are enacted in schools, she drew attention to the ways in which schools "contribute to the reproduction of social relations of inequality between dominant and minority language groups. Using interviews as primary data collection tool, Young (2014) explored how teachers perceived students' home languages and their understanding of plurilingualism. She found that in an effort to meet the needs of plurilingual students, teachers' practice are highly influenced by personal beliefs rooted in monolingual ideologies. Young (2014) proposed that understanding teachers' beliefs/ideologies is a first step in working toward deconstructing and instilling critical language awareness.

Linguicism is defined as "the ideologies, structures, and practices used to legitimate, effectuate, regulate, and reproduce an unequal division of power and resources on the basis of language (their mother tongue" (Skutnabb-Kangas, 1988, p. 13). According to Phillipson (1992), linguicism "has taken over from racism as a more subtle way of hierarchizing social groups in the contemporary world" (p.241). While racism often positions 'whites as the entitled beneficiaries of unearned societal privilege and status' (Huber et al. 2008, 11, 41), linguicism normalizes native (monolingual) speaker values, beliefs, and experiences as those dominant and thus legitimate. Linguicism, particularly a preference toward and the dominance of the English language, is detrimental to the learning experiences of MLLs (Cummins, 2005). An academic prioritization of English frequently results in bilingual children becoming monolingual children (Cummins, 2005). Educators' misunderstandings or lack of knowledge about bi- or multi-lingualism, language learning or acquisition, and cognition are also harmful and exacerbatory (Mellom, et al, 2018). Therefore, disrupting misconceptions about language proficiency is important in teacher education as we strive to prepare teachers to meet the needs of ethnically and linguistically diverse student populations.

Teachers may believe that use of a home language or first language in any context hinders learning a second language (Conteh, 2012; Karabenick & Noda, 2004; Reeves, 2004, 2006). Indeed, educators might even be unaware of why and how home languages can be useful in academic environments (Bailey & Marsden, 2017). In their research study, Karabenick and Noda (2004) found that slightly more than half of all teacher-respondents believed that the use of a first language at home interfered with learning a second language. Mehmedbegovic (2008, 2011) ascertained that educators often had misgivings about the usage of a home language in the classroom setting. This hesitancy was frequently linked with teachers' fears on immigration and difference, or otherness (Burant & Kirby, 2002; Karabenick & Noda, 2004; Mehmedbegovic, 2008, 2011). A tendency of teachers to over-observe cliques of EL students may also reflect a lack of understanding of the other (Karabenick & Noda, 2004). Teachers' lack of prior exposure to diverse populations can manifest as misunderstanding or fear of students (Walker, Shafer & liams, 2004).

Misinformation additionally impacts teachers' negative beliefs. Walker, Shafer and liams (2004) and Reeves (2006) found that teachers believed that MLLs should be fluent in English in either one or two years in classrooms in the United States. However, García's research (2005) showed that most MLLs typically spend six years in the public school system to master the English language. Further issues are the misinterpretation of bilingual education and the overestimation of the associated costs of this type of education (Karabenick & Noda, 2004). In order to meet the needs of increasingly diverse student populations, teachers must go through series of professional development to shift their ideological beliefs that are engrained and overcome their deficit beliefs. There have been numerous studies that showed teacher attitudes can be shifted with appropriate professional development in working with multilingual learners (Fitts & Gross, 2012; Katz, Scott, & Hadjioannou, 2009; Halpern et al., 2022). The authors were asked to work with multiple school districts to create a series of professional development. In an effort to prepare educators for professional development, we set out to determine their pre-existing beliefs about MLLs. It was in that spirit that we adapted a survey to determine the initial beliefs of teachers toward MLLs. The purpose of this study was to determine the beliefs of teachers on multiple factors including culture, cultural competence, parental involvement, and MLLs' struggles in school settings.

Research Question: Our main research question was "What are the perceptions and beliefs of K-12 educators toward their own cultural competence, MLLs' needs and struggles, and parental involvement of MLL families?"

Research Methodology

This study measured variables using a statistical evaluation and analyzes these measurements

using different statistical models in order to create understanding of MLLs' experiences in schooling. The research design of this study provided informative data regarding teachers' beliefs on MLLs' cultural struggles and competence at schools and how that impacts their academic success. The study also provided complete and detailed description of the teachers' perspective on MLLs' difficulties by constructing statistical models and figures to highlight what is observed.

Context of the study

The local context where this study took place included multiple school districts in a state in the Northeast U.S. Participants included both ESL and mainstream teachers in K-12 that taught variety of subjects including ELA, ESL, science, social studies, or mathematics. Convenience sampling and snowball sampling were used to conduct research about participants with specific traits who might otherwise be difficult to identify. Teachers were recruited from the local school districts using electronic communication. The state where this study took place is a small state that have had an increase in the numbers of MLLs in the past ten years. The teacher force consists of predominantly white middle-class women, whom only about 20% of have training to teach MLLs. Therefore, studying the beliefs and perceptions of these teachers is significant.

Instrument

The researchers found Drs. Karabenick and Noda's study in 2004 to be appropriate after they decided to study the beliefs of teachers toward MLLs, and they contacted the authors of the study via email to receive permission for adapting their instrument. After the researchers got permission from Dr. Karabenick and Noda, they start preparing their instrument on Qualtrics software to collect data digitally. The instrument, according to Karabenick and Noda (2004), measures "the teachers' beliefs, attitudes, practices, and needs related to ELLs" (p.56). Karabenick and Noda developed the instrument based on professional literature and many years of teacher experiences. They found that their scale is sufficiently reliable, which had a low but significant correlation (r = .11, p < .001) with teachers' ELL attitudes. The questionnaire reflects MLLs' common difficulties and the impact of parental involvement in schooling.

Karabenick and Noda (2004) used exploratory factor analysis to derive scales (described subsequently), which were constructed by averaging the responses to individual items that had noticeable factor loadings. The items in the instrument were reverse coded where appropriate so that higher values represent more agreement and less disagreement. The instrument included four factors. The first factor labels MLLs' cultural struggles at schools. This factor basically discusses about cultural difficulties that affects language learning. The second factor emphasizes the cultural competence at schools. This factor basically highlights what schools need to do to adjust to diversity. The third factor of the study highlights parental involvement. This factor basically describes what parents need to do to help their children's language learning. The last factor indicates MLLs' common difficulties in classroom. This factor basically highlights common difficulties that ELL students encounter with their classmates and teachers.

Participants

The total sample size of the study was 308 (Male = 121; Female = 187). The occupations of the participants for this study were as follows: ESL/BDL Teachers (168), Special Education Teachers (36), EL Directors (28), School Administrators (28), Assistant Principals (29), and Principals (19). The races of the participants were as follows: White (155), Hispanic (82), Black or African American (45), Asian (26), Alaska Native or American Indian (10), and Native Hawaiian or Other Pacific Islander (6) (see the table 1 below).

Table 1.

Demographics data of the participants (n=308)

		n	% of total
Gender of the respondent	Female	187	60.71
	Male	121	39.29
Occupation	ESOL/BDL Teacher	168	54.55
	Special Education Teacher	36	11.69
	EL Director	28	9.09
	School Administrator	28	9.09
	Assistant Principal	29	9.42
	Principal	19	6.17
Race	White	140	45.45
	Hispanic	81	26.30
	Black or African American	45	14.61
	Asian	26	8.44
	Alaska Native or American Indian	10	3.25
	Native Hawaiian or Other Pacific Islander	6	1.95

The demographic data of this study indicated that 85% of the population hold bachelor's degrees, 11% of the participants hold master's degrees, and 4% of the participants hold doctorate degrees. The demographic data of this study also highlighted that 62% of the participants have more than 10 years teaching experience and 38% of the participants have from 1 to 10 years teaching experience.

Data collection

The researchers had pre-existing professional relationships with many school districts and set out to reach out as many as participants in this state. They contacted school principals, assistant principals, and district coordinators after receiving IRB (#FWA00003132) approval and legally received many teachers' email addresses to send the survey instrument electronically, which was prepared in Qualtrics. The first page of the survey instrument included consent form for the participants. They had to agree to be a participant to go onto further steps of the questionnaire. The electronic link of the instrument was sent to 2203 educators and the researchers only received 308 responses. The participation of the study was completely voluntary. The background of the respondents was not representative of the sample identified by the researchers.

Data analysis

After the researchers collected the raw data, they entered all data into SPSS statistical software to analyze them. AMOS statistical software was used for Confirmatory Factor Analysis (CFA). The researchers basically used descriptive statistics to explain the basic features of the data of the study and provided the samples and the measures done on this study.

Confirmatory Factor Analysis (CFA)

Using AMOS, the researchers conducted Confirmatory Factor Analysis (CFA) to test the instrument. CFA, according to Thompson (2004), "is a multivariate statistical procedure that is used to test how well the measured variables represent the number of constructs" (p.211). The researchers used the CFA to verify the factor structure of a set of observed variables and test the hypothesis that a relationship between observed variables and their underlying latent constructs exists. The result of CFA shows that the instrument supports the statistical data analysis by indicating fix index statistics and recommended value and resources of these statistics.

Table 2 highlighted that the $\chi 2$ (chi-square) ratio with degree of freedom is 523.181/442=1.18. This shows that

Table 2.
Goodness-of-fit indices of the hypothesized measurement model

Fix Index	Resource(s)	Recommended Value	Overall model
CFI	Hu & Bentler (1990); Brown, & Moore (2012)	.90 ≤ CFI ≤ .95 (adequate fit)	.93
RMSEA	Tabachnick & Fidell (2007); Brown, & Moore (2012); Hu & Bentler (1990); Byrne (2004)	RMSEA < .08 (fair fit) RMSEA < .05 (good fit)	.041
SRMR	Hu & Bentler (1999)	SRMR≤.08 (good fit)	.069
TLI	Bentler (1990); Brown, & Moore (2012)	>.95	.98
$\chi 2$ Test of Model Fit	Hu & Bentler (1999); Brown, & Moore (2012); Tabachnick & Fidell (2007); Jöreskog & Sörbom (1993)	Low $\chi 2$ value and p > .05 If p < .05	523.181
$\chi 2$/df	Byrne (2004)	Good Fit $\chi 2$/df < 1 Acceptable Fit $\chi 2$/df < 2	1.183

Note: "$\chi 2$ = chi-square, RMSEA = root mean square error of approximation, SRMR = the standardized root mean square residual, CFI = comparative fit index, and TLI= Tucker-Lewis Index"

the suggested matrix and the original variable matrix are excellent fit. For the measurement model, the value of RMSEA (root mean square error of approximation), which was .041, illustrates good fit for this study. Brown and Moore (2012) affirmed that "the value of RMSEA is good fit if it is <.05 and fair fit if it is <.08" (see also Hu & Bentler, 1999; and Tabachnick & Fidell, 2007) (p.132). According to Hu and Bentler (1999), the value of SRMR (standardized root mean square residual) is an absolute measure of fit and explains "the standardized difference between the observed correlation and the predicted correlation" (p.176). For this study, the researchers found the value of SRMR is .069. Hu and Bentler (1999) affirmed that the value of SRMR should be lower than .08. Brown and Moore pointed out that the value of CFI (comparative fit index) calculates "the model fit by examining the discrepancy between the data and the hypothesized model, while adjusting for the issues of sample size inherent in the chi-squared test of model fit, and the normed fit index" (p.198). For this study, the value of CFI is .93. The value of CFI, according to Brown and Moore (2012), should be between .90 and .95 for a good fit. The value of TLI (Tucker-Lewis Index), which was also called NNFI (Non-Normed Fit Index), was .96. This shows the model of interest improves the fit by 95% relative to the null model. The values of fix index shows that the data fit a hypothesized measurement model and measured variables represent the construct well (see Table 2).

Validity and reliability

The faculty members, who were experts in the field of applied linguistics or closely related field, tested the survey instrument in order to review and revise the instrument before it was administered. They revised unclear terminology, vague items, and inappropriateness by reading each items loudly. The correlation (*rs* = .541, *p* = .001) is a medium/moderate correlation (.40 -.60) (see Tashakkori & Teddlie 2002; Morse, 1994). The validity results of the study were statistically found significant; therefore, this is a valid instrument.

The survey instrument was piloted with 30 participants to evaluate the feasibility, which helped to not waste recourses and time. This helped to test the proposed study process and design. Cronbach's alpha of the study was found to analyze the reliability of scale. The average of 35 items' Cronbach's alpha score was α = .93. Nunnally (1978) affirmed that "a minimum value of .70 for Cronbach's alpha is considered acceptable" (p.54). Hence, the scale was found reliable.

Findings

The researchers used SPSS statistical software to report descriptive statistics regarding teachers' perspective of MLLs' difficulties in the classroom. The findings consisted of four factors as follows: MLLs' cultural struggles at schools, cultural competence at schools, parental involvement, and MLLs' common difficulties in the classroom.

Factor 1: MLLs' cultural struggles at schools

Table 3.

MLLs' cultural struggles at schools

I believe:	Strongly Disagree (%)	Disagree (%)	Neutral (%)	Agree (%)	Strongly Agree (%)	Mean	SD	Cronbach's alpha
Cultural differences enrich the lives of members of communities	5.00	8.10	5.72	58.10	23.08	3.91	1.07	.89
Cultural conflicts arise between ELLs and non-MLLs in class	4.56	6.77	9.29	53.24	26.14	3.89	1.02	.90
ELLs' home culture and native language should be considered in curriculum development	4.37	6.90	10.60	35.10	43.03	4.10	1.06	.94
Language and cultural interventions should be provided in the curriculum	4.10	5.61	9.14	39.70	40.44	4.10	1.10	.94
ELLs' home culture and language should be considered in the special education evaluation process	4.00	5.29	7.32	41.14	42.25	4.21	1.09	.91
Speaking native language at home prevent ELLs from learning English	4.86	8.00	8.86	37.13	41.15	4.12	1.09	.92
Teachers should build a cultural bridge between native and non-native students	5.38	8.31	8.59	39.59	38.14	3.79	1.01	.87
MLLs feel comfortable in the classroom	24.31	19.12	9.00	23.43	24.14	4.01	1.04	.79
Strong native-language skills contribute to MLLs' academic achievement	5.32	7.04	7.91	41.45	38.29	4.20	1.10	.92

Note: "1 = Strongly Disagree, 2 = Disagree, 3 = Neutral, 4 = Agree, and 5 = Strongly Agree."

Table 2 highlighted that cultural differences enhance the lives of members of communities and may cause conflicts between MLLs and non-MLLs in class. The literature also indicated that conflicts can occur between MLLs and non-MLLs because of differences in personality, opinions, and values (Arias-Valenzuela, Amiot, & Ryder, 2019). According to the relatively high mean scores of the item, MLLs' home culture and native language should be considered in curriculum development, and language and cultural interventions should be provided in the curriculum. Curriculum specialist should identify some necessary cultural contents be involved into the curriculum for the students who have culturally and linguistically diverse background (Offorma, 2016). The respondents also mostly agreed that speaking native language at home may prevent MLLs from learning English. Strong native-language skills contribute to MLLs' academic achievement as well. Overall, this factor highlighted that teachers should build a cultural bridge between native and non-native languages for MLLs' academic success. In this factor, cronbach's alpha scores for each item has relatively high (.79<items<.94), which indicated each item under this factor is reliable (see Table 3). For cronbach's alpha, according to Nunnally (1978), "a minimum value of .70 is considered acceptable" (p.243).

Factor 2: Cultural competence at schools

The relatively high mean scores of the items under this factor highlighted that schools should have a set of principles and values that identify diversity and conduct self-assessment to confirm sensitivity to cultural characteristics. Schools need to become culturally competent when there is an issue or crisis, a shared vision, and a desired outcome (Ozfidan & Toprak, 2019). Schools should demonstrate attitudes, behaviors, policies, and structures that allow MLLs to work efficiently cross-culturally and value diversity. According to Sherr and Jones (2019), "cultural competence in social work practice implies a heightened consciousness of how culturally diverse populations experience their uniqueness and deal with their differences and similarities within a larger social context" (p.102). This factor also displayed that schools should be committed to manage the dynamics of difference and incorporate cultural knowledge into their practices and learn about MLLs. Additionally, the findings presented cronbach's alpha scores for each item in this factor. The lowest value of Cronbach's alpha is .87 and the highest of Cronbach's alpha is .95, which indicates each item in the table is reliable (see Table 4).

Factor 3: Parental involvement

This factor highlighted importance of parental involvement in schooling. Most of the respondents emphasized that parents cannot speak English and they only speak their native language at home. According to Ilhan, Ozfidan, and Yilmaz (2019), this is good for parents and children to not forget their native language, but this is not good for parents to learn English and participate in society. The respondents mostly agreed that parents should be welcomed as valuable contributors to school's learning community, and they can support their children's understanding of school's information more effectively when they understand it. Therefore, parents should learn English to help their children in some of the school content. According to the relatively high mean scores of the items, parents should get involved in school activities and they should also monitor and help the completion of their children's homework. Teachers also believed that parents should keep in regular contact with a staff member or teacher regarding their child's progress and visit the school and their child's classroom regularly. Furthermore, the findings also showed that parents should not consider cultural/ethnic differences as a barrier of involving in society. Cronbach's alpha scores for each item in this factor are higher than .85, which indicates each item in the table is reliable (see Table 5).

Factor 4: MLLs' common difficulties in classroom

Table 4.

Cultural Competence at schools

Schools should:	Strongly Disagree (%)	Disagree (%)	Neutral (%)	Agree (%)	Strongly Agree (%)	Mean	SD	Cronbach's alpha
have a set of principles and values that identify diversity	6.10	10.10	10.00	52.72	21.08	3.79	1.01	.87
conduct self-assessment to confirm sensitivity to cultural characteristics	5.56	9.24	9.25	49.77	26.18	4.01	1.05	.91
demonstrate attitudes, behaviors, "policies, and structures that allow them to work efficiently cross-culturally and value diversity	6.30	10.10	10.60	36.10	36.90	3.92	1.00	.89
be committed to manage the dynamics of difference	4.61	7.10	9.70	42.10	35.48	4.02	1.05	.92
adjust to diversity and the communities' cultural contexts	5.00	8.30	9.31	38.14	39.25	4.10	1.07	.93
incorporate cultural knowledge into their practices and" learn about	6.86	9.00	8.86	36.13	39.15	4.12	1.09	.95

Note: "1 = Strongly Disagree, 2 = Disagree, 3 = Neutral, 4 = Agree, and 5 = Strongly Agree."

Table 5.

Parental involvement

Parents:	Strongly Disagree (%)	Disagree (%)	Neutral (%)	Agree (%)	Strongly Agree (%)	Mean	SD	Cronbach's alpha
cannot speak English	6.72	10.10	9.05	41.10	33.03	4.01	1.03	.91
speak their native language at home	5.56	8.78	10.28	49.24	26.14	4.05	1.09	.92
should be welcomed as valuable contributors to school's learning community	4.37	7.90	9.60	36.10	42.03	4.09	1.06	.94
can support their children's understanding of school's information more effectively when they understand it	5.10	5.61	9.14	38.70	40.44	4.11	1.10	.91
should get involved in school activities	4.29	5.00	7.32	41.14	42.25	4.13	1.09	.94
should monitor and help the completion of their children's homework	4.13	7.00	9.86	37.86	41.15	4.12	1.09	.87
should keep in regular contact with a staff member or teacher regarding their child's progress	5.35	6.25	7.20	38.97	42.23	4.06	1.05	.92
should visit the school and their child's classroom regularly	5.31	8.38	8.59	39.59	38.14	3.83	1.01	.85
should not consider cultural/ ethnic differences as a barrier of involving the society	5.25	6.35	8.20	37.97	42.23	4.13	1.10	.93

Note: "1 = Strongly Disagree, 2 = Disagree, 3 = Neutral, 4 = Agree, and 5 = Strongly Agree."

Table 6.

ELLs' common difficulties in classroom

MLLs:	Strongly Disagree (%)	Disagree (%)	Neutral (%)	Agree (%)	Strongly Agree (%)	Mean	SD	Cronbach's alpha
should develop literacy in their first language to facilitate the development of writing and reading in English.	4.72	7.10	9.00	53.10	24.08	4.14	1.01	.93
are able to use their native language to access academic content more easily	4.77	7.56	9.29	42.24	36.14	4.09	1.04	.91
do not understand the content compared to native English speakers	5.37	6.90	9.60	36.10	42.03	4.10	1.06	.94
will be more successful if they learn to write and read in their native language	4.61	6.10	10.14	39.70	38.44	4.11	1.10	.94
exposed to English to learn better	5.00	5.29	7.32	40.14	42.25	4.13	1.07	.91
are more successful after they solved the language barrier	4.86	8.00	8.86	37.13	41.15	4.12	1.09	.92
don't take their study seriously	4.25	6.35	8.20	38.97	42.23	4.20	1.05	.93
speak more of their native language than English	5.38	6.31	8.59	39.59	40.14	3.79	1.01	.87
become too dependent on the teacher	4.31	9.12	9.00	38.43	39.14	4.01	1.04	.90
encounter culture shock	4.32	6.04	7.91	42.45	39.29	4.21	1.10	.92
are bored and/or unmotivated	5.56	6.77	9.29	52.24	26.14	4.23	1.02	.94

Note: "1 = Strongly Disagree, 2 = Disagree, 3 = Neutral, 4 = Agree, and 5 = Strongly Agree."

This factor highlighted MLLs' common difficulties in the classroom. According to the relatively high mean scores of the items, MLLs should develop literacy in their first language to facilitate the development of writing and reading in English. This is one of the biggest issues for people who have poor language background from their native language. MLLs do not understand the content compared to native English speakers. This affects MLLs' academic achievement negatively. The findings indicate that MLLs are able to use their native language to access academic content more easily and they will be more successful if they learn to write and read in their native language. MLLs encounter culture shock; therefore, they are usually bored and unmotivated. Teachers should find more integrative and entertainment teaching strategies for MLLs, who have motivation problems and are bored in classroom (Gowri & Ilankumaran, 2020). For reliability, the researchers reported Cronbach's alpha scores, which were relatively high (.87<items<.94) (see Table 6).

Discussion

The purpose of this study was to explore the beliefs and perceptions of educators toward MLLs. The findings of the study presented mixed results. The results indicated that although teachers' beliefs about MLLs are slightly negative their overall beliefs are positive. Many teachers while welcoming MLLs in their classrooms, desiring to support them, are unsure of how to do so. This is critical in the field of TESOL/ bi-/multilingual education in discussing the effect of beliefs and attitudes on teaching. This finding is consistent with other teacher beliefs studies that beliefs could implicitly influence the pedagogical choices that teachers make on a daily basis (Harrison & Lakin, 2018). As Fives and Buehl (2012) indicate teachers' implicit beliefs are critical determinant of teacher behaviors and "act as a filter on interpretation of teaching experiences without the teacher's conscious knowledge" (Harrison & Lakin, 2018, pg. 97). Although the majority of the teachers in this study expressed a positive attitude in teaching MLLs, their responses are consistent with previous literature that there is a dire need for professional development to support these teachers in order to ensure MLLs receive the best education possible (Karabenick & Noda, 2004).

The researchers adapted Karabenick and Noda's (2004) survey to measure teachers' beliefs regarding their own cultural competence, multilingual learners' needs, and parental involvement. Karabenick and Noda considered it important to measure teachers' successed goals, owing to their current dominance in the conceptual framework on school motivation. This method differentiates among teachers' focus on mastery versus performance. Mastery goals highlight involvement in individual improvement and tasks. On the contrary, performance goals highlight students' abilities, which are made more noticeable comparing inter-student. The current study found that MLLs are more likely than non- MLLs to suffer from interpersonal comparisons and competition because of language and cultural differences. Thus, MLLs are more likely to thrive in classrooms that are more mastery focused and less performance focused. The current study investigated whether teachers with more positive attitudes toward MLLs stressed the more beneficial goals for achievement.

The results of this study indicated that teachers believed that speaking native language at home prevents MLLs from learning English, and MLLs' home culture and native language should be considered in curriculum development and language and cultural interventions should be provided in the curriculum. Teachers' beliefs also indicated that strong native-language skills contribute to MLLs' academic achievement and teachers should build a cultural bridge between native and non-native students. The beliefs of teachers indicated that schools should have a set of principles and values that identify diversity and conduct self-assessment to confirm sensitivity to cultural characteristics. The results also highlighted that parents should be welcomed as valuable contributors to school's learning community, and they can support their children's understanding of school's information more effectively when they understand it. The teachers beliefs revealed that they expected parents to be involved in school activities and they should also monitor and help the completion of their children's homework. Teachers expected parents to keep regular contact with a staff member or teacher regarding their child's progress and visit the school and their child's classroom regularly. Even though the teachers believed that parents should not consider cultural/ethnic differences as a barrier of involving the society, their beliefs on parental involvement certainly was not one that was culturally sensitive to different parental backgrounds, similar to what Gonzales and Gabel (2017) highlighted about the different ways parents can be involved despite what is expected by the U.S. schools. The teachers belief about parental involvement was in line with western views of parent involvement. The teachers also believed in the power of first language literacy to facilitate the development of writing and reading in English. MLLs do not understand the content compared to native English speakers. Additionally, teachers believed that MLLs are able to use their native language to access academic content more easily and they will be more successful if they learn to write and read in their native language.

Conclusion

Understanding teachers' beliefs is an important step in understanding teaching practices and their overall

impact on MLL achievement in schools. By exploring teachers' beliefs and attitudes we can expand our understanding of teacher practices and design professional learning experiences that delve deeper on self-awareness of unconscious attitudes toward MLLs. MLLs are the fastest growing student population in the United States (NCES, 2016), and will comprise nearly 40% of the K-12 population by 2030 (Center for Public Education, 2012). There is an increased need for teachers of MLLs to be more prepared to teach MLLs and gain greater understanding of the needs of these students and approach teaching MLLs with a more additive approach rather than deficit views. More research on teachers' beliefs about MLLs is required and its impact on everyday practice should also be further investigated. Further studies should include qualitative measures that may integrate teacher interviews and classroom observations to see how self-reported measures of teachers correlate with everyday attitudes toward MLLs.

Overall, the researchers suggested that teachers of MLLs need to be made aware of the instructional practices with the support of continuing discussion and professional learning to apply foundational knowledge into practice. The current research provided insights into teachers' everyday beliefs and instructional practices in working with MLLs; however, more research is needed on teacher beliefs on teaching and learning to further understand the instructional needs of teachers of MLLs. The current research not only provided teachers a space for reflection on their beliefs and practices in the classroom, but it also informs school leaders on the importance of making bottom-up decisions regarding professional learning needs of the teachers who are already demonstrated their commitment to supporting all of their students but specifically MLLs with more culturally responsive instruction. For a future study, the researchers are planning to include qualitative measures that might integrate teacher interviews and classroom observations to see how self-reported measures of teachers correlate with everyday instructional practices with MLLs.

References

Abdi, H., & Asadi, B. (2015). A Synopsis of Researches on Teachers' and Students' Beliefs about Language Learning. International Journal on Studies in English Language and Literature (IJSELL), 3(4), 104-114.

Arias-Valenzuela, M., Amiot, C. E., & Ryder, A. G. (2019). Identity configurations and well-being during normative cultural conflict: The roles of multiculturals' conflict management strategies and academic stage. *European Journal of Social Psychology*, *49*(5), 970-991

Bailey, E. G., & Marsden, E. (2017). Teachers' views on recognizing and using home languages in predominantly monolingual primary schools. *Language and Education, 31*(4).

Balderrama, M. V. (2001). The (mis) preparation of teachers in the proposition 227 era: Humanizing teacher roles and their practice. *The Urban Review, 33*(3), 255-267.

Barton, A., Bragg, J., & Serratrice, L. (2009). "Discovering language" in primary school: An evaluation of a language awareness programme. *Language Learning Journal,* 37(2), 145-164.

Batlova, J., & McHugh, M. (2010). *Number and growth of students in US schools in need of English instruction.* Washington, DC: Migration Policy Institute.

Bentler, P.M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin*, 107, 238-246.

Borg, S. (2003). Teacher cognition in language teaching: A review of research on what language teachers think, know, believe, and do. *Language Teaching*, 36, 81–109.

Brown, T.A., & Moore, M.T. (2012). Confirmatory factor analysis. *Handbook of structural equation modeling*, 361-379.

Byrne, B. M. (2004). Testing for multigroup invariance using AMOS graphics: A road less traveled. *Structural Equation Modeling*, *11*(2), 272-300.

Burant, T. J., & Kirby, D. (2002). Beyond classroom-based early field experiences: Understanding an "educative practicum" in an urban school and community. *Teaching and Teacher Education, 18*(5), 561-575.

Butcher, J., Sinka, I., & Troman, G. (2007). Exploring diversity: Teacher education policy and bilingualism. *Research Papers in Education, 22*(4), 483-501.

Center for Public Education (2012). *The United States of education: The changing demographics of the United States and their schools*. Retrieved from: http://www.centerforpubliceducation.org/You-May-Also-Be-Interested-In-landing-page-level/Organizing-a-School-YMABI/The-United-States-of-education-The-changing-demographics-of-the-United-States-and-their-schools.html

Conteh, J. (2003). *Succeeding in diversity: Culture, language and learning in primary classrooms* Stoke-on-Trent, England: Trentham Books.

Cummins, J. (2005). A proposal for action: Strategies for recognising heritage language competence as a learning resource within the mainstream classroom. *The Modern Language Journal, 89*(4), 585-591.

Fang, Z. 1996. A review of research on teacher beliefs and practices. *Educational Research*, 38: 47–65.

Fenstermacher, G. D. 1994. "The knower and the known: The nature of knowledge in research on teaching". In *Review of research in education*, Edited by: Darling-Hammond, L. Vol. 20, 3–56. Washington, DC: American Educational Research Association.

Fives, H., & Buehl, M. M. (2012). Spring cleaning for the "messy" construct of teachers' beliefs: What are they? Which have been examined? What can they tell us?. In K. R. Harris, S. Graham, & T. Urdan (Eds.), *APA educational psychology handbook* (Vol. 2)(pp. 471–479). Washington, DC, USA: American Psychological Association.

Fitts, S., Gross, L. A. (2012). Teacher candidates learning from English learners: Constructing concepts of language and culture in Tuesday's tutors after-school program. *Teacher Education Quarterly, 39, 75-95.*

Garcia, E. (2005). Teaching and learning in two languages. *Bilingualism and schooling in the United States*. New York: Teachers College Press.

Gowri, K. J., & Ilankumaran, M. (2020). Strategies in Teaching English at The Tertiary Level of Education. *Tathapi with ISSN 2320-0693 is an UGC CARE Journal*, *19*(44), 295-302.

Halpern, C., Ozfidan, B., & Rasool, S. (2022). Pre-Service Teachers' Perceptions of Cultural Competence to Teach Culturally and Linguistically Diverse Students. *Journal of multilingual and multicultural development*, 1-16.

Haney, J., Czerniak, C., & Lumpe, A. (1996). Teacher Beliefs and Intentions Regarding the Implementation of Science Education Reform Strands. Journal of Research in Science Teaching, 33(9), 971-993.

Harrison, J. & Lakin, J (2018). Mainstream Teachers' Implicit Beliefs about English Language Learners: An Implicit Association Test Study of Teacher Beliefs. *Journal of Language, Identity, and Education*, 17:2, 85-102.

Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, *6*(1), 1-55.

Huber, L. P., C. B. Lopez, M. Malagon, V. Velez, and D. G. Solorzano. 2008. "Getting beyond the 'Symptom', Acknowledging the 'Disease': Theorizing Racist Nativism." *Contemporary Justice Review: Issues in Criminal, Social, and Restorative Justice* 11 (1): 39–51.10.1080/10282580701850397

Ilhan, F., Ozfidan, B., & Yilmaz, S. (2019). Home Visit Effectiveness on Students' Classroom Behavior and Academic Achievement. *Journal of Social Studies Education Research*, *10*(1), 61-80.

Jöreskog, K. G., & Sörbom, D. (1993). *LISREL 8: Structural equation modeling with the S I M P L I S command language*. Scientific Software International.

Kagan, D. M. 1992. Implications of research on teacher belief. *Educational Psychologist*, 27: 65–90.

Karabenick, S. A., & Noda, P. A. C. (2004). Professional development implications of teachers' beliefs and attitudes toward English language learners. *Bilingual Research Journal*, *28*(1), 55-75.

Katz, L., Scott, J.C., & Hadjioannou, X. Exploring Attitudes toward Language Differences: Implications for Teacher Education Programs. (2009). *Affirming Students' Right to Their Own Language: Bridging Educational Policies and Pedagogical Practices* (pp.99-116). New York, NY: Routledge/ Urbana, IL: NCTE.

Kang, H. J. (2022). Preservice Elementary Teachers' Understanding of Fraction Multiplication and Division in Multiple Contexts. *International Electronic Journal of Elementary Education*, *15*(2), 109-121.

Kelly, L. B. (2018) Preservice Teachers' Developing Conceptions of Teaching English Learners. *TESOL Quarterly, 52:1,* 110-136.

Khader, F. R. (2012). Teachers' Pedagogical Beliefs and Actual Classroom Practices in Social Studies Instruction. *American International Journal of Contemporary Research*, 2(1), 73-92.

Knopp, T., & Smith, R. (2005). A brief historical context for dispositions in teacher education. In R. Smith, D. Skarbek, & J. Hurst (Eds.), *The passion of teaching: Dispositions in the schools* (pp. 1–13). Lanham, MD: Scarecrow Education.

Lippi-Green, R. (2012). *English with an Accent: Language, Ideology and Discrimination in the United States*. New York: Routledge.

Mantero, M., & McVicker, P. (2006). The impact of experience and coursework: Perceptions of second language learners in the mainstream classroom. *Radical Pedagogy, 8*(1) Retrieved from http://radicalpedagogy.icaap.org/content/issue8_1/mantero.html

Martin-Jones, M. (2009). Bilingualism, education and the regulation of access to language resources. In M. Heller (Ed.), *Bilingualism: A social approach* (pp. 161–182). London: Palgrave Macmillan.

McSwain, A. (2001). The effects of multicultural and bilingual training on preservice students' self-reported level of competency. *Multiple Voices, 5*, 54–65.

Mehmedbegovic, D. (2008). Leading increasingly linguistically diverse london schools. *Educate, 8*, 4-21.

Mehmedbegovic, D. (2011). *A study in attitudes to languages in England and Wales*. Saarbruken, Germany: LAP Lambert Academic Publishing.

Mellom, P. J., Straubhaar, R., Balderas, C., Ariail, M., & Portes, P. R. (2018). "They come with nothing:" How professional development in a culturally responsive pedagogy shapes teacher attitudes towards Latino/a English language learners. *Teaching and Teacher Education, 71,* 98-107.

Midgley, C., Ed. (2002). Goals, goal structures, and patterns of adaptive learning. Mahwah, NJ: Erlbaum.

Midgley, C., Kaplan, A., & Middleton, M. (2001). Performance-approach goals: Good for what, for whom, under what circumstances, and at what cost? Journal of Educational Psychology, 93, 77–86.

Morse, J. M. (1994). *Critical issues in qualitative research methods*. Thousand Oaks, CA: Sage Publications Inc.

National Center for Education Statistics (2019). *English Language Learners in Public Schools.* Retrieved from https://nces.ed.gov/programs/coe/indicator_cgf.asp

Nicholls, J. (1984). Achievement motivation: Conceptions of ability, subjective experience, task choice, and performance. Psychological Review, 91, 328–346.

Nunnally, J. C. (1978). *Psychometric theory* (2nd ed.). New York: McGraw-Hill.

Offorma, G. C. (2016). Integrating Components of Culture in Curriculum Planning. *International Journal of Curriculum and Instruction*, *8*(1), 1-8.

Ozfidan, B. (2017). Academic Value of Bilingual Education: Factors in Learning on Culture and Language. *Revista de Cercetare şi Intervenţie Socială*, (59), 34-47.

Ozfidan, B., & Toprak, M. (2019). Cultural awareness on a bilingual education: A mixed method study. *Multicultural Learning and Teaching*, *15*(1).

Ozfidan, B., & El-Dakhs, D. A. S. (2023). Students' perceptions of online learning in non-L1 environments. *International journal of multilingualism*, 1-15.

Pajares, F. (1992). Teachers' beliefs and educational research: Cleaning up a messy construct. *Review of Educational Research*, 62, 307–332.

Phillips, S., & Borg, S. (2009). Exploring Tensions between Teachers' Grammar Teaching Beliefs and Practices. *System*, 37(2009), 380-390.

Phillipson, R. (1992). *Linguistic imperialism and linguicism, Linguistic Imperialism*. Oxford: OUP, pp. 50-57.

Reeves, J. (2004). "Like everybody else": Equalizing educational opportunity for English language learners. *TESOL Quarterly, 38*, 43-66.

Reeves, J. (2006). Secondary teacher attitudes toward including English-language learners in mainstream classrooms. *Journal of Educational Research, 99*, 131-142.

Richardson, V. (1996). The Role of Attitudes and Beliefs in Learning to Teach. In J. Sikula, T. J. Buttery, & Guyton, E. (Eds.), Handbook of Research in Teacher Education, Second Edition (pp. 102-119). Macmillan, New York.

Rueda, R. and Garcia, E. 1994. *Teachers' beliefs about reading assessment with Latino language minority students*, Washington, DC: Office of Educational Research and Improvement. Retrieved from ERIC database.

Sherr, M. E., & Jones, J. M. (2019). *Introduction to competence-based social work: The profession of caring, knowing, and serving*. Oxford University Press.

Skutnabb-Kangas, T. (1988) "Multilingualism and the education of minority children." *Minority Education: From Shame to Struggle.* 9-44.

Sugimoto, A. T., Carter, K., and Stoehr, K. J. (2017). Teaching "in their best interest": Preservice teachers' narratives regarding English Learners. *Teaching and Teacher Education,* 67, 179-188.

Tabachnick, B. G. & Fidel, L. S. (2007). *Using multivariate statistics*, 5th (Edn.). Pearson Publication.

Tashakkori, A. & Teddlie, C. (Eds.) (2002). *Handbook of mixed methods for the social and behavioral sciences.* Thousand Oaks, CA: SAGE Publications.

Thompson, A. (1992). Teachers' beliefs and conceptions. In D. Grouws (Ed.), *Handbook of research on mathematics teaching and learning* (pp. 127–146). New York, NY: Macmillan.

Thompson, B. (2004). *Exploratory and confirmatory factor analysis: Understanding concepts and applications*. American Psychological Association.

Ukpokodu, O. N. (2007). Preparing socially conscious teachers: A social justice oriented teacher education. *Multicultural Education*, 15(1), 8–15.

Valdes, G., Kibler, A., & Walqui, A. (2014). *Changes in the expertise of ESL professionals: Knowledge and action in an era of new standards*. Alexandria, VA: TESOL International Association.

Walker, A., Shafer, J., & Iiams, M. (2004). "Not in my classroom": Teacher attitudes towards English language learners in the mainstream classroom. *National Association for Bilingual Education Journal of Research and Practice*, 2, 130– 160.

Wright, W. (2006). A catch-22 for language learners. *Educational Leadership*, 64(3), 22–27

Young, A. S. (2014). Unpacking teachers' language ideologies: attitudes, beliefs, and practiced language policies in schools in Alsace, France. *Language Awareness, 23:1-2,* 157-171.

APPENDIX A

Demographic Information

Which one of the following categories best describes your job position or primary role?

☐ ESL/BDL Teacher
☐ Teacher (Please specify grade span) Early Childhood Elementary Secondary
☐ EL Director
☐ Special Education Teacher
☐ School Administrator
☐ Guidance Counselor
☐ School Psychologist
☐ Social Worker
☐ School Support Staff (e.g. paraprofessional, classroom aid)
☐ Administrative Support (e.g. registrar, attendance secretary, etc.)
☐ Other, please specify:

Are you certified in ESL/BDL? ☐ Yes ☐ No

Have you received any training on ELs?

☐ Yes
☐ No
☐ Other, please specify

Has your training on ELs helped you better support all of your students (both ELs and non-ELs)?

☐ Yes, it has.
☐ No, but it has helped me better support ELs.
☐ No, the training was not helpful at all.
☐ Not applicable.

What is your race? (Select one or more):

☐ Black or African American
☐ Alaska Native or American Indian
☐ Asian
☐ Native Hawaiian or Other Pacific Islander
☐ White
☐ Other
☐ Decline to State

What is your gender?

☐ Female
☐ Male
☐ Decline to State

What is your highest level of education?

☐ Bachelors
☐ Masters
☐ Doctorate
☐ Other

How long have you worked in the field of education?

☐ Less than 1 year
☐ 1 - 3 years
☐ 4 - 6 years
☐ 7 - 10 years
☐ More than 10 years

How many years have you been at your current position?

☐ Less than 1 year
☐ 1 - 3 years
☐ 4 - 6 years
☐ 7 - 10 years
☐ More than 10 years

Have you ever had ELs in your classes?

☐ Yes
☐ No
☐ N/A

Do you have ELs in your classes now?

☐ Yes
☐ No
☐ N/A

Likert-scale question

The different sections of the survey asks for the extent to which you agree or disagree with a series of statements. To ensure complete coverage, and to increase the survey's reliability, there may be more than one statement that refers to the same domain. Please indicate your response to each statement.

I believe:	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
Cultural differences enrich the lives of members of communities					
Cultural conflicts arise between ELLs and non-MLLs in class					
ELLs' home culture and native language should be considered in curriculum development					
Language and cultural interventions should be provided in the curriculum					
ELLs' home culture and language should be considered in the special education evaluation process					
Speaking native language at home prevent ELLs from learning English					
Teachers should build a cultural bridge between native and non-native students					
MLLs feel comfortable in the classroom					
Strong native-language skills contribute to MLLs' academic achievement					

Please continue providing your degree of agreement or disagreement with each statement...

Schools should:	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
have a set of principles and values that identify diversity					
conduct self-assessment to confirm sensitivity to cultural characteristics					
demonstrate attitudes, behaviors, "policies, and structures that allow them to work efficiently cross-culturally and value diversity					
be committed to manage the dynamics of difference					
adjust to diversity and the communities' cultural contexts					
incorporate cultural knowledge into their practices and" learn about					

Please continue providing your degree of agreement or disagreement with each statement...

Parents:	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
cannot speak English					
speak their native language at home					
should be welcomed as valuable contributors to school's learning community					
can support their children's understanding of school's information more effectively when they understand it					
should get involved in school activities					
should monitor and help the completion of their children's homework					
should keep in regular contact with a staff member or teacher regarding their child's progress					
should visit the school and their child's class-room regularly					
should not consider cultural/ethnic differences as a barrier of involving the society					

Please continue providing your degree of agreement or disagreement with each statement...

MLLs:	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
should develop literacy in their first language to facilitate the development of writing and reading in English.					
are able to use their native language to access academic content more easily					
do not understand the content compared to native English speakers					
will be more successful if they learn to write and read in their native language					
exposed to English to learn better					
are more successful after they solved the lan-guage barrier					
don't take their study seriously					
speak more of their native language than English					
become too dependent on the teacher					
encounter culture shock					
are bored and/or unmotivated					

This page is intentionally left blank.

www.iejee.com

iejee

International Electronic Journal of Elementary Education
September 2023, Volume 16, Issue 1, 75-88

A Bibliometric Review of Research on Education for Sustainable Development, 2019-2023

Mira Esti Kusumaningrum[a,*], Heru Kuswanto[b], Jumadi[c], Slamet Suyanto[d], Suhartini[e], Desy Purwasih[f], Ratna Prabawati[g]

Received : 23 April 2023
Revised : 10 July 2023
Accepted : 29 September 2023
DOI : 10.26822/iejee.2023.315

[a] **Corresponding Author:** Mira Esti Kusumaningrum, Science Education,Yogyakarta State University, Indonesia.
E-mail: miraesti.2022@student.uny.ac.id
ORCID: https://orcid.org/0009-0005-2908-7722

[b] Heru Kuswanto, Science Education,Yogyakarta State University, Indonesia.
E-mail: herukus61@uny.ac.id
ORCID: https://orcid.org/0000-0002-2693-8078

[c] Jumadi, Science Education,Yogyakarta State University, Indonesia.
E-mail: jumadi@uny.ac.id
ORCID: https://orcid.org/0000-0002-4055-5065

[d] Slamet Suyanto, Science Education,Yogyakarta State University, Indonesia.
E-mail: slamet_suyanto@uny.ac.id
ORCID: https://orcid.org/0000-0002-9581-0596

[e] Suhartini, Science Education,Yogyakarta State University, Indonesia.
E-mail: suhartini@uny.ac.id

[f] Desy Purwasih, Science Education,Yogyakarta State University, Indonesia.
E-mail: desypurwasih.2022@student.uny.ac.id
ORCID: https://orcid.org/0000-0002-4258-2059

[g] Ratna Prabawati, Science Education,Yogyakarta State University, Indonesia.
E-mail: ratnaprabawati.2022@student.uny.ac.id
ORCID: https://orcid.org/0009-0009-4834-069X

Abstract

The purpose of this study is to identify current research trends in the area of ESD. The SLR (Systematic Literature Review) method was used in this bibliometric investigation. Sourced from Scopus from 2019 to 2023, the database is used. Research on Education for Sustainable Development (ESD) has seen a significantly increased attention in recent years. Therefore, it is necessary to conduct a detailed examination regarding this matter promptly. This can help enhance our ability to identify study trends and better understand the character of the ESD method. Using a bibliometric approach, we identified 3,247 journal articles that containing the keyword "education for sustainable development". The results of the study show that scientific publications about ESD have relatively increased. After filtering with certain criteria, 60 articles were obtained. Consequently, Sweden contributed the most documents on ESD and the Futur Institute from Germany predominated in publications. The most contributing author is Gericle n. The research trends on ESD were grouped into five clusters, namely 1) ESD in the scope of learning; 2) ESD on educational programs and policies; 3) ESD linked to learning and teaching in practice; 4) ESD related to achievements or competencies in continuous learning; and 5) ESD related to continuous evaluation. Research findings can assist researchers to understand the ESD research trends and provide recommendations for further research directions

Keywords:

Education for Sustainable Development, Sustainable Development, Bibliometric, Review

Introduction

During the 1990s, educators were worried about the waste of resources and the negative impact of economic development on the environment, leading to an increased emphasis on environmental education (Jickling, B., & Wals, A. E., 2012). This marked the early stages of a concept that later known as "Education for Sustainable Development" (Aikens, K., et al., 2018; UNESCO, 2005; Jickling, B., & Wals, A. E., 2012). The international policy community has been raising awareness and integrating education into global policy initiatives with sustainable development goals, and these efforts have increased in the last two decades (Jickling, B.,

& Wals, A. E., 2012). An example of the efforts was the launch of the United Nations Decade of Education for Sustainable Development in 2004, to integrate the values, principles, and practices of sustainable development into formal and informal education (UNESCO, 2005).

Citizens of every nation need the education to pave the way for the development of knowledge and attitudes to create a sustainable society (Aikens, K., et al., 2018; Salas-Zapata, W.A., et al, 2018). The United Nations adopted 17 sustainable development goals (SDGs) reflecting the expansion of international efforts to create a sustainable society at the end of the 2015 Education Decade for Sustainable Development. Hallinger, P., & Chatpinyakoop, C. (2019) conclude that education is able to foster values, attitudes, and sustainable behavior among the next generation of global citizens, which is the key to achieving all of the SDGs, as well as providing the transformation required for sustainable nation (Kioupi, V., & Voulvoulis, N., 2019).

UNESCO defines Education for Sustainable Development (ESD) as a bridge for learners of all ages, knowledge, skills, values, and agency to address the interrelated global challenges of climate change, loss of biodiversity, unsustainable use of resources, and inequality (UNESCO, 2023) and empower communities to take responsibility for creating a sustainable future (UNESCO, 2003). ESD is an approach that combines behavior change, educational pedagogy, and knowledge continuity where the integration of these three things results in an effective and transformative learning process by empowering students to make informed decisions and take responsible actions for environmental integrity, economic feasibility, and just society for present and future generations while respecting cultural diversity (Redman, E., & Larson, K., 2011; Taimur, S., & Sattar, H., 2020).

Education is the most strategic way to instill and apply the values of sustainable development. In addition, education is considered the most strategic in encouraging sustainable development and increasing human capacity to be able to overcome environmental and development issues and problems. ESD is a multidisciplinary concept that views the concept of development from a social, economic, and environmental perspective (Vilmala, B. K., et al, 2022). The aim of ESD is to develop competencies that enable and empower individuals to reflect on their own actions by considering their current and future social, cultural, economic, and environmental impacts from a local and global perspective, taking part and being responsible for creating a sustainable community, and developing skills, values, and attitudes that enable people (including students) to lead healthy lives and be able to respond to local and global challenges (UNESCO, 2018).

Wals, A. E., & Lenglet, F. (2016) state that by the end of the 20th century, society will contribute to building a sustainable personality. Society is required to understand the complex world in which people live and understand how to deal with uncertainty, risk, and the high speed of social change. Communities must also be able to collaborate, communicate, and act positively in responding to world changes. To meet these demands, competence is required. UNESCO (2018) has defined eight main ESD competencies, namely 1) systemic thinking competence, 2) anticipatory competence, 3) normative competence, 4) strategic competence, 5) collaborative competence, 6) critical thinking competence, 7) self-awareness competence, and 8) integrated problem-solving competence. Each competency has its own qualities and areas of relevance. The eight competencies are interdependent and need to be developed basic competencies. In addition, basic competencies such as communication skills are very important to deal with sustainable development. (Wiek et al., 2011).

The importance of Education for Sustainable Development (ESD) is motivated by the alarming state of the Earth, such as floods, landslides, hurricanes, and global warming. Therefore, in addressing these issues, all parties need to contribute, including in the field of education with predetermined competencies. The potential of education as an effective means of solving problems on earth can be realized if the education system adheres to sustainable development (UNESCO, 2018).

Previous bibliometric studies were conducted to analyze the literature relating to sustainability and education in higher education settings. The topics analyzed included health, education, management, energy, agriculture, and environmental issues. The results showed that the largest publications were written in the US, UK, China, Australia, and Canada (Veiga Avila, L., et al., 2018). Similar research was continued by Hallinger, P., & Chatpinyakoop, C. (2019) with literature from 1998 to 2018, and the results of the author's co-citation analysis revealed three research groups that underlaid this knowledge base, namely Managing Sustainability in Higher Education, HESD Competence, and HESD implementation. Previous studies examined the HESD knowledge base from the perspective of implementation barriers (Lozano, R., et al., 2015; Velazquez, L., et al., 2005), teacher education (Chinedu, C. C., et al., 2018; Velazquez, L., et al., 2006), teaching and learning (Bostrom, M., et al., 2018; Ferreira, J.A., 2009; Rickinson, M., et al., 2008; Stanislas, M., et al., 2018), and conceptual models (Brebler, J. & Kappler, S., 2017; Scott, W., 2015).

According to this higher education research, Prieto-Jiménez, E., et al. (2021) concluded that the SDGs in general, SDG 4 regarding quality education, education

for sustainable development, higher education, and education management were the five core clusters identified in his bibliometric mapping analysis. In conclusion, the necessity of altering the purpose and role of higher education is emphasized more in order to address sustainable development.

According to relevant research on bibliometric review studies from ESD by Grosseck, G., et al. (2019), which retrieved 1813 papers from 1992 to 2018 using bibliometrics, these studies illustrate two main research directions for the entire time period, which involve integrating sustainable development into education and education into sustainable development. Moreover, research on Education for Sustainable Development (ESD) has expanded quickly, as seen by the rise in publications, authors, and journals. This is consistent with research findings from Hallinger, P., & Nguyen, V. T. (2020), which found that publishing rates have been rising quickly and that understanding of ESD has risen significantly over the previous 30 years. The bibliometric study methodology is further employed by Hallinger, P., et al. (2020) in additional research, this time related to simulations and serious games used in sustainability education from 1997 to 2019. This study provides a general summary of future research that will be conducted to support the effects of simulations and serious games on learner attitudes, knowledge, and behavior.

From the previous works, we can see that ESD may be examined from various fields. As a result, this study aims to continue and uncover the latest trends in ESD research. While earlier studies covered a wider range of topics, this study is more narrowly focused on schools and its surroundings. The bibliometric review of this study aims to expand on previous research reviews by mapping education for sustainable development from 2019 to 2023 with the required criteria. This review provides a benchmark for future research reviews on ESD, revealing emerging research trends from the interdisciplinary field. The research questions to guide the review of this study are as follows.

1) What is the output of publications, source documents, and language on ESD in the years 2019-2023?

2) What is the extent of the publication about ESD across countries and institutions in the world?

3) Who are the top authors researching ESD in the world?

4) What is the pattern of publication about ESD based on journals?

5) What are the results of the visualization of research trends on ESD?

Research Method

This study used bibliometric analysis and was expected to provide valuable references for future research. Briner, R., & Denyer, D., (2012) state that SLR plays a role in systematically collecting literature and providing descriptive analysis such as journal publications, trends, creating insights, to developing a base of knowledge on any topic. The steps of the SLR process followed were descriptions of Govindan & Hasanagic (2018); Mathiyazhagan et al. (2021) for this study with some modifications as needed. The database optimizes the Scopus database used in the initial phase to collect relevant articles. As Scopus uses consistent criteria to select papers for inclusion in its index, the author decided to use it as the document source for this evaluation. It also includes a greater selection of Web of Science papers to review social science and education research (Hallinger, P., et al., 2019; Mongeon, P., et al., 2016). Furthermore, Scopus provides more complex options for exporting bibliographic data than Google Scholar. The phrase "education for sustainable development" was used as a filter to search for titles, abstracts, and keywords from 2019 to 2023. Data collection was carried out in February-March 2023. A total of 3,247 documents met the search criteria. However, in the second stage, further screening was carried out to meet what was expected, such as open access documents, certain subject areas, and documents taken only from journals in English. Based on these provisions, 60 new documents were obtained. Later, the spreadsheet application program was used to analyze the data. A deeper analysis was conducted to analyze research trends, including publication output characteristics, document sources, distribution of countries and institutions, distribution of outcomes in the subject category, top authors, top citations, and publication trends from 2019 to 2023. The application or software used is VoSViewer to find out research trends on ESD. More detailed stages can be seen in Figure 1.

Findings and Discussion

Number of ESD documents in the 2019-2023

There are 3,247 articles related to ESD research in the Scopus database. All articles found were filtered based on the use of language, namely English. Then, the articles underwent a further screening process with certain criteria, such as All Open Access to facilitate access during cross-checking of data, keywords related to ESD and education, type of article document, and affiliations of eight institutions and top 15 countries, resulting in the amount of new data as many as 60 articles. The ESD research graph for the years 2019-2023 with these certain criteria can be seen in Figure 2.

Figure 1.
Summary of SLR of ESD stage with modification

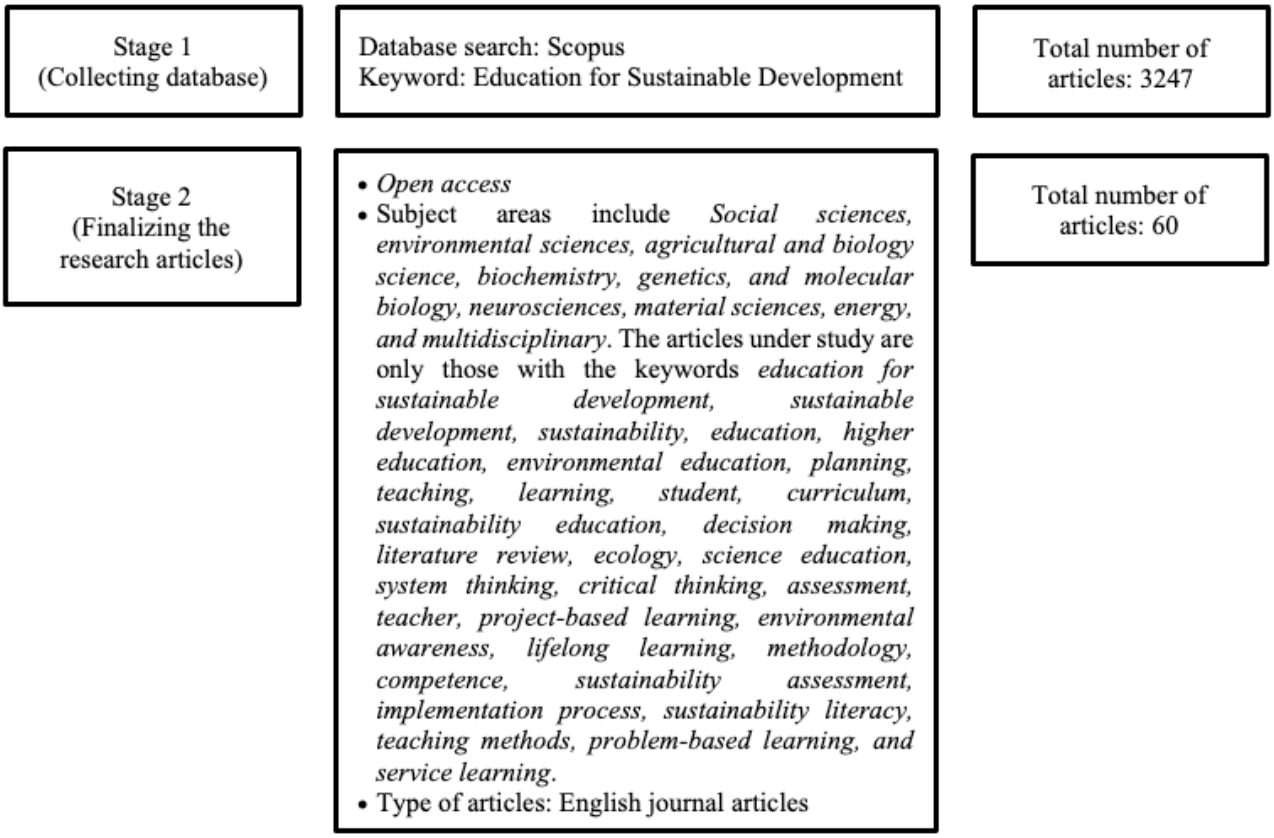
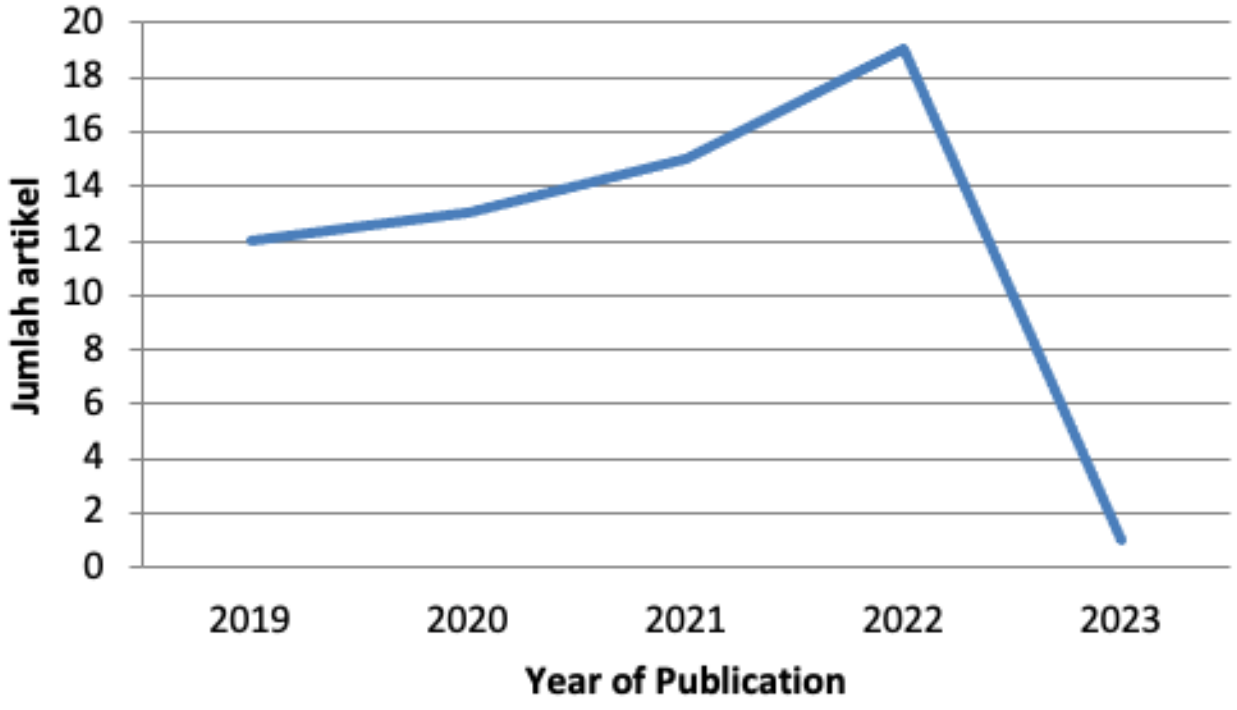


Figure 2.
Number of ESD articles in 2019-2023



The number of publications from 2019 to 2022 is depicted in Figure 2 and shows an increase. However, the year 2023 indicates a decrease due to only one article being discovered. This is reasonable considering that 2023 had only been operational for three months at the time this study. Figure 2 displays the outcomes of the increased number of ESD documents. This is consistent with the findings of Hallinger, P., & Chatpinyakoop, C. (2019), which found that from 1998 to 2018 there was a growing tendency, despite seven years of marginally lower growth.

3247 documents were discovered after developing a search based on the concept of "Education for Sustainable Development" (searches continued until March 2023). 60 documents meeting a predetermined set of criteria were obtained. This study can be seen as a follow-up to earlier research conducted by other researchers. The majority of these studies, however, focus on particular aspects of sustainable development, while the focus of this research is ESD in education and all things educational. The annual increase in ESD research is depicted in Figure 2. This is achievable because, between 2000 and 2015, significant efforts were made to assist and improve the fundamental requirements of the world's poorest nations, one of which is education.

The best approach to establish morals and social norms in children, both formally and informally, is through education. As a result, education is highly essential in developing attitudes that support sustainable development. One of the most significant aspects affecting environmentally conscious behavior is education. The correlation between environmental awareness and knowledge demonstrates the value of education. According to research by Iswari, R. D., and Utomo, S. W. (2017), an individual's concern for environmental issues is positively connected with their knowledge of nature and the environment. It is hardly unexpected that ESD research is growing every year considering its significance.

Basically, the number of publications has been steadily increasing since the concept of ESD was officially recognized internationally. This shows that ESD is still an interesting research area. According to Wright, T., and Pullen, S., (2007), the creation and adoption of the 2030 Agenda for sustainable development along with the 17 SDGs by all United Nations member countries in 2015, provide a simple explanation of this growth. As a result, papers published after 2015 place greater emphasis on the role of education in achieving sustainable development.

Distribution of the countries of publication

For the distribution of 60 articles based on country of origin, in terms of collaboration between authors, 11 countries were selected that had ESD documents with a minimum of two documents produced, a number of citations, and strength of collaboration between authors. The detailed information provided in Table 1.

Table 1 shows that Sweden has the highest number of documents, namely 23 documents. This outcome was attributed to the publication of ESD policy documents by UNESCO (UNESCO 2004, 2014, 2017) which have had an impact on education policies around the world including Sweden's national guidance document for education (SNAE, 2013; Skolverket, 2011). Gericke, N., & Torbjörnsson, T. (2022) concluded that all types of schools that provide strategic priority to ESD in Sweden implement ESD, including in other countries such as Portugal (Spinola, H., 2015) and Belgium (Boeve-de Pauw, J., & Van Petegem, P. (2011).

Additionally, there was an increasing emphasis in Sweden on including ESD as one of the fundamental concepts of education. Swedish schools were making an effort to collaborate with ESD. ESD is an integral component of academic work rather than an exclusive approach to other school activities. The use of projects to assist student learning about sustainable development is a crucial strategy. The commitment of the teachers and administrators who support the work is a requirement for the success of the schools, all of which have a particular reputation in the field of ESD (Fredriksson, U., N., et al., 2020).

Table 1.
Distribution of articles about ESD in each country

Country	Documents	Citation		Total link strength
		Total citation	Citation per article	
Sweden	23	200	8.7	14
Germany	20	175	8.8	7
Spanish	14	169	12.1	1
Belgium	3	81	27	5
Norway	3	8	2.7	5
Portugal	3	36	12	5
Dutch	3	18	6	3
Taiwan	2	41	20.5	4
South Africa	2	14	7	2
United States	2	28	14	2
Indonesia	3	1	0.3	0

On the other hand, each nation has a unique cultural heritage and level of development, which might significantly influence their understanding of ESD (Wang, W., 2015). OECD member nations dominate the list of nations with the most publications and citations. This can be due to the fact that developing nations do not fund ESD-related research. The number of universities and research institutions in OECD member countries is higher than in non-OECD member countries, which may be contributing to the sharp rise in publications and citations from these nations. Wright, T., and Pullen, S., (2007) claimed that although bibliometric data may not unequivocally support this claim, it remains a factor worthy of consideration in future studies.

Next, the eleven countries involved in writing ESD then have a relationship that can be visualized in Figure 3.

Figure 3.
Relationship of the distribution of article origin

Indonesia has only one citation and does not have collaboration between authors and other countries, so it can be seen in Figure 3 that Indonesia was separated from other countries. In contrast to Sweden, which has been obliged to implement ESD, policies and implementation of ESD in Indonesia still need to be determined on a national planning scale (Suprastowo, 2010) even though Bappenas (National Development Planning Bureau) has compiled a National Medium-Term Development Plan document which contains all of the SDGs agenda within the national policy framework (Ministry of Education and Culture, 2021). Meanwhile, the interrelated relationships of the 10 countries in ESD research can be seen in Figure 4.

Figure 4.
Collaboration among authors in each country

Figure 4 shows the interconnections among 10 countries. Sweden has a larger circle and more connecting lines indicating that it has the most documents and author collaborations. Figure 4 displays the most significant nations in the field as huge circles, together with the degree of communication between nations as cooperating nodes. Figure 4 demonstrates how a number of nations frequently combine to form a single, sizable cluster, with an average of one or two core countries in each cluster. Three frequently mentioned countries in articles and citations were: Sweden, Germany, and Spain. On the other hand, other core nations include South Africa, Taiwan, Belgium, Norway, Portugal, Netherlands, and the United States. Due to the fact that ESD has drawn the interest of academics from all over the world, we can draw the conclusion that "geographical advantage is not the main factor influencing cooperative relations" (Liao, H., et al., 2018).

There are 157 organizations inside institutions that have published on ESD, according to cross-organizational results. The final six organizations are listed in Table 2 after additional selection was made based on at least two papers from each organization. The majority of research was conducted by institutions from Sweden, Germany, and Spain, the three largest countries, as shown in Table 2. In terms of the research activities carried out by each institution, the findings can be summarized into the following categories: There was no institutional coordination, each university from the three major countries tends to conduct its own research, and each work receives a significant number of citations. This is conceivable because the research priorities of various countries may differ.

Table 2.
Number of ESD articles (2019-2023) of inter-organization

Organization	Documents	Citation	Total link strength
Department of Computer Architecture, Universitat Politècnica De Catalun-ya—Upc-Barcelonatech, Barcelona, 08034, Spain	2	7	0
Department of Environ-mental and Life Sciences, Karlstad University, Karlstad, Sweden	2	6	0
Department of Pedagogical, Curricular and Professional Studies, University of Gothen-burg, Gothenburg, Sweden	2	14	0
Institut Futur, Freie Universität Berlin, Berlin, 14195, Germany	4	55	0
Karlstad University, Karlstad, Sweden	3	73	0
university of gothenburg, goteborg, sweden	2	0	0

Top Authors in ESD Research

In the data processing related to the collaboration of the authors, a total of six authors related to ESD were produced. Among the six, three of them have the same collaborative power. This can be seen in Figure 5.

Figure 5.
Relationship of inter-author collaboration

The explanatory data from Figure 5 can be seen in Table 3 regarding the number of documents, citations, and the strength of collaboration for each author.

Table 3.
Documents, Citation, and Top Collaboration Strength

Author	Documents	Citation		Total link strength
		Total citation	Citation per document	
Gericle n.	12	145	12.1	10
Olsson d.	6	100	16.7	10
Boeve-de pauw j.	5	99	19.8	10
Barth m.	5	39	7.8	0
Singer-brow-dowski m.	5	42	8.4	0
Sanchez carracedo f.	9	121	13.4	0

Figure 5 shows the collaborative relationship between authors, and it can be seen that the red relationship lines form a single bond with the names Gericle n., Olsson d., and Boeve-de pauw j. This indicates that the three of them have collaborated in writing about ESD and is supported by the data in Table 3, that the three have a relationship strength of up to 10. The "bubbles" in Figure 5 represent authors. The lines represent working relationships, whereas the nodes represent paper usage. The three authors of One Line have worked together on one or more publications.

In the map, collaborating authors frequently cluster together while non-collaborating authors are spread out. Basic issues like "Who collaborates with whom and how?" "Can past collaborations be assessed?" "What happened to collaborations between industry and universities?" etc. are typically the focus of cooperation analysis. Along with the quantity of coauthored works (publications), we should also consider "the extent to which co-authorship becomes international in the sense of collaboration" (Ellegaard, O., & Wallin, J. A., 2015) when describing the academic quality (impact) of a researcher.

ESD publication patterns based on journals

Table 4 illustrates the journals that have made the most significant contributions to ESD research, with the top five journals displayed. Sustainability Switzerland has the highest number of publications, with 23 articles. This is possible because the scope of the journal relates to sustainability. It can also be seen in Table 4 that the journals in which these publications appear are indexed in Scopus with Quartile Q1 and Q2.

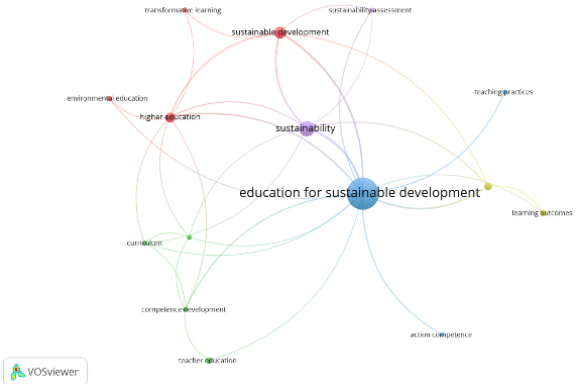
As a result, ESD has garnered considerable interest among academics, particularly since many of its studies have been published in highly regarded international publications. According to the quartiles of these journals, there were 23 publications in Q2 and 17 in Q1, indicating that researchers generally adopt a modest publication strategy. The majority of research in this field displays good overall quality. The majority of the references pertain to environmental education and public health, thereby contributing to the dissemination of journals that help in assessing how different disciplines have an impact on the ESD domain.

Table 4.
Number of ESD Documents in Each Journal

Journal	No. of Documents	Scopus Quartile
Sustainability Switzerland	23	Q2
Environmental Education Research	7	Q1
Journal of Environmental Education	5	Q1
Journal of Cleaner Production	3	Q1
International Journal of Environmental Research and Public Health	2	Q1

Trends of ESD Research in 2019-2023

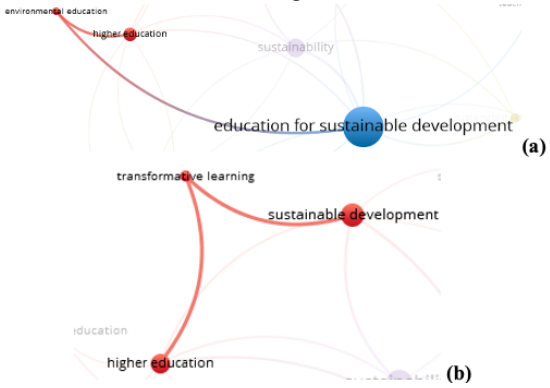
Figure 6.
Overall Profile of ESD Research from 2019 to 2023



Among the 60 articles related to ESD research in the Scopus database, research trends on this topic can be visualized using the VoSViewer software. This effort helps find research novelty. The findings show that there were several parameters or interrelationships between variables in ESD, such as research on education for sustainable development, sustainability, sustainable development, higher education, environmental education, ESD with higher education, ESD on competencies, and curriculum policies. Figure 6 shows a visualization of the overall research on ESD. Researchers around the world produced five clusters.

The first cluster was ESD in the learning environment. The second cluster was ESD on education programs and policies. The third cluster was ESD which is related to learning and teaching in practice. The fourth cluster shows ESD related to achievements or competencies in continuous learning, and cluster five was ESD related to continuous evaluation. Nearly every word or subject in each cluster has a connection to the phrase ESD. The main subject field in which the research was conducted and the potential issues with ESD are both indicated by the fact that each word has at least one relationship to the others. The large blue ESD circle with lines connecting it to other keywords may be seen in Figure 6. This demonstrates the significance of ESD as a study issue.

Figure 7.
ESD in relation to teaching



In Cluster 1, as shown in Figure 7, ESD was closely related to higher education, transformative learning, sustainable development, and environmental education. However, upon closer examination based on the keywords, higher education, transformative learning, and sustainable development were not directly related to ESD. It is different from environmental education which is related to higher education, both of which are directly related to ESD. Even so, the four keywords can be made into one cluster on the basis of the scope of learning.

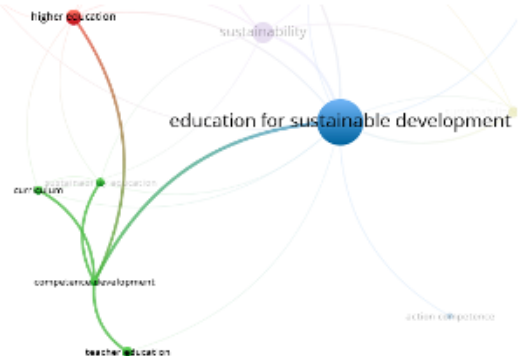
Figure 7a demonstrates how ESD and environmental education were related. Environmental education in the narrow sense is essentially promoted by some UNESCO programs that use limited environmental interpretations; other programs interpret the environment more broadly and considering social factors; still other programs use ESD frameworks or incorporate ESD themes and approaches to reframe existing curricula and still refer to them as EE. The quality envisioned by transformational education should instead be used as the benchmark for evaluating EE and ESD rather than only one another (Plavova, M., 2013).

Next, figure 7b shows how sustainable development was connected to higher education and transformative learning. As it addresses common research topics like comprehending ESD definitions, concepts, curricula, or theories, exploring global research agendas and practices on ESD, mainstreaming ESD into educational policies, etc., sustainable development is a crucial topic throughout the period (Grosseck, G., et al., 2019). Transformative learning, on the other hand, can support ESD because its goal is to encourage critical self-reflection, which results in a shift in perspective and a change in behavior. Our obligation to ensure that these changes in perspective and behavior are rooted in restorative principles is at the core of the saguf message and program (Balsiger, J., et al.).

The global sustainability agenda has a substantial impact on higher education for sustainable development (HEfSD). Many higher education institutions actively working to incorporate the Sustainable Development Goals (SDGs) into HEfSD policies, curricula, and practices through sporadic and isolated initiatives. These institutions are responsible for preparing the next generation of sustainability leaders with the necessary knowledge and skills.

The success of HEfSD in policy, curriculum, and practice is largely contingent on a deeper understanding of what is already in place, according to research findings (Franco, I., et al., 2019). Gaps, focus areas, parallels, and divergences exist among regional HEfSD agendas. In order to successfully integrate HEfSD into policies, curricula, and practices that align with the SDGs and the overarching goals of the Global Agenda for Sustainable Development, it is expected to provide both conceptual and practical tools to higher education institutions and stakeholders across regions.

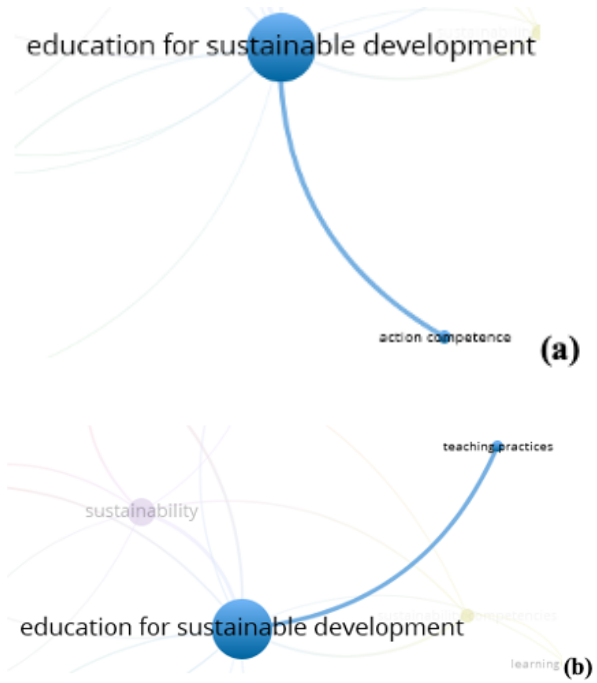
Figure 8.
ESD in relation to educational policy programs



Furthermore, curriculum, sustainable education, competence development, and teacher education are grouped into one cluster and are closely related to ESD. Figure 8 shows that the four keywords are also related to higher education. This is possible because more research has been conducted at the higher education level. Thus, the four keywords are grouped into one cluster on the basis of education programs and policies.

Considering the United Nations Decade for ESD (2004-2015), competencies are heavily explored in the context of cross-curricular issues like Sustainable Development and Education for Sustainable Development (ESD). Competency development appears to be more of a continuous learning process, hence opportunities for systematic and comprehensive competency development must be offered not just once, but also in subsequent courses throughout the entire education of future teachers at universities and other settings. In order to assist the development of ESD-specific abilities in teacher education learning formats that facilitate authentic and task-oriented contact, real-world problems, and the challenges of engaging partners in practice are also required (Brandt, J. O., et al., 2019).

Figure 9.
ESD in relation to teaching and learning practices



Different from clusters one and two, in cluster 3 there are two keywords, namely 'action competence' and 'teaching practices'. Even though both are in one cluster, it turns out that the two were not related to each other. However, each was directly related to ESD. This can be seen in Figure 9. Action competence and teaching practices were grouped into one cluster on the basis of learning and teaching in practice.

According to Mogensen and Schneack (2010), action competence—or "Bildung" in German—is the guiding principle of education, and the idea of "action" serves as the foundation for action competence. When examining ESD from an action competency perspective, education should come first. Furthermore, this strategy interprets "competence" considerably differently than that associated with the individualistic paradigm of human resource management. The OEC and the DeSeCo (Definition and Selection of Competency) perspective, which Development promotes, as well as the subject-focused ideas of competency, have some similarities and distinctions, though.

The action competency approach's quality standards are as follows: they emphasize better teaching and learning, reflect the democratic values that ESD aspires to advance, collaborate with key stakeholders, and support institutional and individual learning. As a result, this strategy offers an illustration from a pedagogical standpoint.

Specialized training is required for ESD-focused instruction, particularly to comprehend the ESD idea itself. Even though most teachers do not comprehend ESD, they acknowledge employing participatory approaches when teaching, according to a study (Anyolo, E. O., et al., 2018). However, in the real world, professors exclusively employ lectures, question-and-answer sessions, presentations, and lectures. Although this cannot be applied generally, it does highlight how crucial it is for teachers to socialize ESD. Therefore, research on ESD and teaching practices is gaining momentum and becoming a global issue.

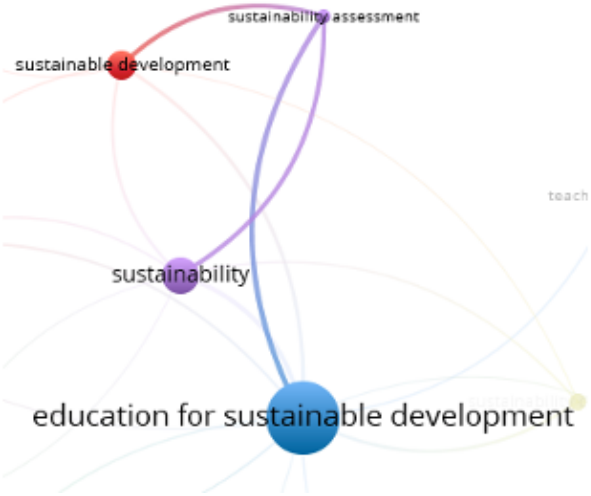
Figure 10.
ESD in relation to competency outcomes

Figure 10 presents the relationship in cluster four which consists of learning outcomes and sustainability competencies. Unlike keywords in cluster three, the keywords in cluster four were related to each other and were directly related to ESD. Both were classified in one cluster on the basis of achievement or competence in sustainable teaching.

"Education for Sustainable Development" appears to have had a big influence on learning results. According to the United Nations Conference on Environment and Development (1992), one of the objectives of Agenda 21 is to alter people's attitudes and character in addition to creating values, attitudes, and behaviors that are compatible with sustainable

development. Significant affective outcomes are part of the "education for sustainable development" goal, which higher education institutions find challenging to meet. This group also includes learning outcomes because they are a fascinating subject. It is expected that each student would develop the attitudes, feelings, and motivations required to shape their conduct and allow them to become environmentally conscious "actors" in line with the goal of "education for sustainable development."

Figure 11.
ESD in relation to the evaluation



Finally, Figure 11 presents the relationship between the keywords 'sustainability' and 'sustainable assessment'. Both of these keywords were directly related to ESD. However, it can be seen that the sustainability assessment was also related to sustainable development. It can be concluded that cluster five was grouped based on the sustainable evaluation.

The majority of indicators and assessment questions in education were related to the curriculum and the availability of courses on sustainability, but Yarime, M. and Tanaka, Y. (2012) found that sustainability is more than just a topic to be added to an already overloaded curriculum. Additionally, sustainability can provide access to various viewpoints on organizational change, teaching, policy, and particularly the philosophy of continuous assessment.

In conclusion, several studies on ESD have been conducted. However, several studies focus on, among others, instrument development (Widodo, A., et al., 2023), exploration of sustainable development competencies in certain disciplines (Fröberg, A., et al., 2022), pilot studies (Eichinger, M ., et al., 2022), socialization of TPACK-ESD (Purwianingsih, W., et al., 2022), analysis of the process of integration of ESD in various fields of education (Singer-Brodowski, M., et al., 2019), ESD in higher education (Singer-Brodowski, M., et al., 2022; Birdman, J., et al., 2022; Heinrichs, H., 2021; Weiss, M., et al., 2021; Busquets, P., et al., 2021; Sánchez-Carracedo, F., & López, D., 2021; Sánchez-Carracedo, F., et al., 2021; Sánchez-Carracedo, F., et al., 2021; Finnveden, G., et al., 2020; Muñoz-Rodríguez, J. M., et al., 2020; Duarte, A. J., et al., 2019; Westermark, Å., & Jansund, B., 2019; Sundermann, A., & Fischer, D., 2019).

This diversity led to the current research being conducted to determine the mapping of the latest research trends in ESD through bibliometric studies and to develop the research of Hallinger, P., & Chatpinyakoop, C. (2019) with literature from 1998 to 2018 that focused solely on higher education. In simple terms, in the last five years (2019-2023), the number of ESD documents throughout the year has experienced a significant increase and was divided into five clusters.

Conclusion

Based on the findings of this analysis, it can be stated that the 2019–2023 ESD research trend is likely to continue to rise. This aligns with earlier research' findings from 1990 to 2018, which similarly tended to show a sharp rise. Due to its history of using ESD in education, Sweden has produced the most study up to present time. The following organization or institution that contributes the most documents to publications on ESD is the future institute from Germany. Gericle is the top author on ESD according to Scopus data. The trend of research on ESD produces five large clusters, namely 1) ESD in the scope of learning; 2) ESD on educational programs and policies; 3) ESD related to learning and teaching in practice; 4) ESD related to achievements or competencies in continuous learning; and 5) ESD related to sustainable evaluation.

ESD will continue to be a hot topic in international study until early 2023. Thus, by considering the clusters identified in this study, research on ESD can be further enhanced. ESD research can contribute to achieving sustainable development objectives, particularly in the area of education. A sense of concern and of responsibility for the environment can be developed as well as remedies to the problems associated with global issues by improving the number and quality of ESD research.

References

Aikens, K., McKenzie, M., & Vaughter, P. (2018). Environmental and sustainability education policy research: A systematic review of methodological and thematic trends. Environmental and Sustainability Education Policy, 265-292.

Anyolo, E. O., Kärkkäinen, S., & Keinonen, T. (2018). Implementing education for sustainable development in Namibia: School teachers' perceptions and teaching practices. *v*(1), 64-81.

Balsiger, J., Förster, R., Mader, C., Nagel, U., Sironi, H., Wilhelm, S., & Zimmermann, A. B. (2017). Transformative learning and education for sustainable development. *GAIA-ecological Perspectives for Science and Society, 26*(4), 357-359.

Birdman, J., Barth, M., & Lang, D. (2022). Connecting curricula and competence through student learning journeys. *Sustainability: Science, Practice and Policy, 18*(1), 560-575.

Brandt, J. O., Bürgener, L., Barth, M., & Redman, A. (2019). Becoming a competent teacher in education for sustainable development: Learning outcomes and processes in teacher education. *International Journal of Sustainability in Higher Education, 20*(4), 630-653.

Breßler, J.; Kappler, S. (2017). A Systematic Review of Education for Sustainable Development (No. 007); Chemnitz Economic Papers; Chemnitz University of Technology, Faculty of Economics and Business Administration: Chemnitz, Germany.

Briner, R., Denyer, D., 2012. Systematic Review and Evidence Synthesis as a Practice and Scholarship Tool. In: Handbook of evidence-based management: Companies, classrooms and research, pp. 112–129. https://doi.org/10.1093/oxfordhb/9780199763986.013.0007.

Boeve-de Pauw, J., & Van Petegem, P. (2011). The effect of Flemish eco-schools on student environmental knowledge, attitudes, and affect. *International Journal of Science Education, 33*(11), 1513-1538.

Boström, M., Andersson, E., Berg, M., Gustafsson, K., Gustavsson, E., Hysing, E., ... & Öhman, J. (2018). Conditions for transformative learning for sustainable development: A theoretical review and approach. *Sustainability, 10*(12), 4479.

Busquets, P., Segalas, J., Gomera, A., Antúnez, M., Ruiz-Morales, J., Albareda-Tiana, S., & Miñano, R. (2021). Sustainability education in the Spanish higher education system: Faculty practice, concerns and needs. *Sustainability, 13*(15), 8389.

Chinedu, C. C., Wan-Mohamed, W. A., & Ogbonnia, A. A. (2018). A systematic review on education for sustainable development: Enhancing TVE teacher training programme. *Journal of Technical Education and Training, 10*(1).

Duarte, A. J., Malheiro, B., Arnó, E., Perat, I., Silva, M. F., Fuentes-Durá, P., ... & Ferreira, P. (2019). Engineering education for sustainable development: the European project semester approach. *IEEE Transactions on Education, 63*(2), 108-117.

Eichinger, M., Bechtoldt, M., Bui, I. T. M., Grund, J., Keller, J., Lau, A. G., ... & Heinzel, S. (2022). Evaluating the Public Climate School—A School-Based Programme to Promote Climate Awareness and Action in Students: Protocol of a Cluster-Controlled Pilot Study. *International Journal of Environmental Research and Public Health, 19*(13), 8039.

Ellegaard, O., & Wallin, J. A. (2015). The bibliometric analysis of scholarly production: How great is the impact?. Scientometrics, 105, 1809-1831.

Ferreira, J. A. (2018). Unsettling orthodoxies: Education for the environment/for sustainability. In Environmental and Sustainability Education Policy (pp. 121-134). Routledge.

Finnveden, G., Friman, E., Mogren, A., Palmer, H., Sund, P., Carstedt, G., ... & Svärd, L. (2020). Evaluation of integration of sustainable development in higher education in Sweden. *International Journal of Sustainability in Higher Education, 21*(4), 685-698.

Franco, I., Saito, O., Vaughter, P., Whereat, J., Kanie, N., & Takemoto, K. (2019). Higher education for sustainable development: Actioning the global goals in policy, curriculum and practice. *Sustainability Science, 14,* 1621-1642.

Fredriksson, U., N. Kusanagi, K., Gougoulakis, P., Matsuda, Y., & Kitamura, Y. (2020). A comparative study of curriculums for education for sustainable development (ESD) in Sweden and Japan. *Sustainability, 12*(3), 1123.

Fröberg, A., Wiklander, P., & Lundvall, S. (2022). Sustainable Development Competencies among More than 1100 Certified Physical Education and Health Teachers in Sweden. *International Journal of Environmental Research and Public Health, 19*(23), 15914.

Govindan, K., Hasanagic, M., 2018. A systematic review on drivers, barriers, and practices towards circular economy: a supply chain perspective. Int. J. Prod. Res. 56 (1–2), 278–311. https://doi.org/10.1080/00207543.2017.1402141.

Grosseck, G., Țîru, L. G., & Bran, R. A. (2019). Education for sustainable development: Evolution and perspectives: A bibliometric review of research, 1992–2018. *Sustainability, 11*(21), 6136.

Hallinger, P., & Chatpinyakoop, C. (2019). A bibliometric review of research on higher education for sustainable development, 1998–2018. *Sustainability, 11*(8), 2401.

Hallinger, P., & Nguyen, V. T. (2020). Mapping the landscape and structure of research on education for sustainable development: A bibliometric review. *Sustainability, 12*(5), 1947.

Hallinger, P., Wang, R., Chatpinyakoop, C., Nguyen, V. T., & Nguyen, U. P. (2020). A bibliometric review of research on simulations and serious games used in educating for sustainability, 1997–2019. Journal of Cleaner Production, 256, 120358.

Heinrichs, H. (2021). Teaching Sustainable Development in a Sensory and Artful Way—Concepts, Methods, and Examples. *Sustainability, 13*(24), 13619.

Iswari, R. D., & Utomo, S. W. (2017). Evaluasi penerapan program adiwiyata untuk membentuk perilaku peduli lingkungan di kalangan siswa (Kasus: SMA Negeri 9 Tangerang Selatan dan MA Negeri 1 Serpong). *Jurnal Ilmu Lingkungan, 15*(1), 35-41.

Jickling, B., & Wals, A. E. (2012). Debating education for sustainable development 20 years after Rio: A conversation between Bob Jickling and Arjen Wals. *Journal of Education for Sustainable Development, 6*(1), 49-57.

Kemdikbud. (2021). Lokakarya Nasional Inisiatif Indonesia Menuju Pendidikan Berkelanjutan (ESD) tahun 2030. https://www.kemdikbud.go.id/main/blog/2021/04/lokakarya- nasional-inisiatif-indonesia-menuju-pendidikan-berkelanjutan-esd-tahun-2030. Diakses 18 April 2023.

Kioupi, V., & Voulvoulis, N. (2019). Education for sustainable development: A systemic framework for connecting the SDGs to educational outcomes. *Sustainability, 11*(21), 6104.

Liao, H., Tang, M., Luo, L., Li, C., Chiclana, F., & Zeng, X. J. (2018). A bibliometric analysis and visualization of medical big data research. *Sustainability, 10*(1), 166.

Lozano, R., Ceulemans, K., Alonso-Almeida, M., Huisingh, D., Lozano, F. J., Waas, T., ... & Hugé, J. (2015). A review of commitment and implementation of sustainable development in higher education: results from a worldwide survey. *Journal of cleaner production,* 108, 1-18.

Mathiyazhagan, K., Rajak, S., Sampurna Panigrahi, S., Agarwal, V., Manani, D., 2021. Reverse supply chain management in manufacturing industry: a systematic review. Int. J. Prod. Perform. *Manage. 70* (4), 859–892. https://doi.org/10.1108/IJPPM-06- 2019-0293

Mogensen, F., & Schnack, K. (2010). The action competence approach and the 'new'discourses of education for sustainable development, competence and quality criteria. *Environmental education research, 16*(1), 59-74.

Mongeon, P., & Paul-Hus, A. (2016). The journal coverage of Web of Science and Scopus: a comparative analysis. *Scientometrics, 106*, 213-228.

Muñoz-Rodríguez, J. M., Sánchez-Carracedo, F., Barrón-Ruiz, Á., & Serrate-González, S. (2020). Are we training in sustainability in higher education? Case study: Education degrees at the University of Salamanca. *Sustainability, 12*(11), 4421.

Pavlova, M. (2013). Towards using transformative education as a benchmark for clarifying differences and similarities between environmental education and education for sustainable development. *Environmental Education Research, 19*(5), 656-672.

Prieto-Jiménez, E., López-Catalán, L., López-Catalán, B., & Domínguez-Fernández, G. (2021). Sustainable development goals and education: A bibliometric mapping analysis. *Sustainability, 13*(4), 2126.

Purwianingsih, W., Novidsa, I., & Riandi, R. (2022). Program for Integrating Education for Sustainable Development (ESD) into Prospective Biology Teachers' Technological Pedagogical Content Knowledge (TPACK). *Jurnal Pendidikan IPA Indonesia, 11*(2).

Redman, E., & Larson, K. (2011). Educating for sustainability: Competencies & practices for transformative action.

Rickinson, M., & Lundholm, C. (2008). Exploring students' learning challenges in environmental education. *Cambridge Journal of Education, 38*(3), 341-353.

Salas-Zapata, W. A., Ríos-Osorio, L. A., & Cardona-Arias, J. A. (2018). Knowledge, attitudes and practices of sustainability: Systematic review 1990–2016. *Journal of Teacher Education for Sustainability, 20*(1), 46-63.

Sánchez-Carracedo, F., & López, D. (2021). A Service-Learning Based Computers Reuse Program. *Sustainability, 13*(14), 7785.

Sánchez-Carracedo, F., Moreno-Pino, F. M., Romero-Portillo, D., & Sureda, B. (2021). Education for sustainable development in Spanish university education degrees. *Sustainability, 13*(3), 1467.

Sánchez-Carracedo, F., Ruiz-Morales, J., Valderrama-Hernández, R., Muñoz-Rodríguez, J. M., & Gomera, A. (2021). Analysis of the presence of sustainability in Higher Education Degrees of the Spanish university system. *Studies in Higher Education, 46*(2), 300-317.

Scott, W. (2014). Education for sustainable development (ESD): A critical review of concept, potential and risk. Schooling for sustainable development in Europe: Concepts, policies and educational experiences at the end of the UN Decade of Education for Sustainable Development, 47-70.

Singer-Brodowski, M., Etzkorn, N., & Von Seggern, J. (2019). One transformation path does not fit all—insights into the diffusion processes of education for sustainable development in different educational areas in Germany. *Sustainability, 11*(1), 269.

Singer-Brodowski, M., Förster, R., Eschenbacher, S., Biberhofer, P., & Getzin, S. (2022). Facing crises of unsustainability: Creating and holding safe enough spaces for transformative learning in higher education for sustainable development. In Frontiers in Education (p. 81). Frontiers.

Skolverket (2011b). Curriculum, Compulsory school, Preschool class, Recreation centre, In English. Skolverket. https://www.skolverket.se/publikationer?id=2687

Spínola, H. (2015). Environmental literacy comparison between students taught in Eco-schools and ordinary schools in the Madeira Island region of Portugal. *Science Education International,* 26, 392-413.

Stanitsas, M., Vareilles, É., Kirytopoulos, K., & Aldanondo, M. (2018, June). Sustainable development in serious games: rethinking game-based learning strategies for master's degree engineers. In MOSIM'18-12ème Conférence internationale de Modélisation, Optimisation et SIMulation (pp. 8-p).

Sundermann, A., & Fischer, D. (2019). How does sustainability become professionally relevant? Exploring the role of sustainability conceptions in first year students. *Sustainability, 11*(19), 5155.

Suprastowo, P. (2010). Kebijakan dan Implementasi Pendidikan untuk Pembangunan Berkelanjutan (Education for Sustainable Development/ESD). Kebijakan dan Implementasi Pendidikan untuk Pembangunan Berkelanjutan (Education for Sustainable Development/ESD).

Swedish National Agency for Education. (2013). Curriculum for the upper secondary school.

Taimur, S., & Sattar, H. (2020). Education for sustainable development and critical thinking competency. *Quality education*, 238-248.

UNESCO. (2023). What you need to know about education for sustainable development. https://www.unesco.org/en/education-sustainable-development/need-know. Diakses 17 April 2023.

UNESCO. (2003). Conférence générale. December 200

UNESCO. United Nations Decade of Education for Sustainable Development (2005–2014). International Implementation Scheme. 2005. Available online: http://unesdoc.unesco.org/images/0014/001486/148654e.pdf

United Nations Educational, Scientific and Cultural Organization (UNESCO). (2014). UNESCO roadmap for implementing the global action programme on education for sustainable development. https://unesdoc.unesco.org/ark:/48223/pf0000230514

United Nations Educational, Scientific and Cultural Organization (UNESCO). (2017). Education for sustainable development goals: Learning objectives. http://unesdoc.unesco.org/images/0024/002474/247444e.pdf

United Nations Conference on Environment and Development. 1992. Agenda 21 Chapter 36. Accessed July 5, 2023. http://www.un-documents.net/a21-36.htm

UNESCO. (2004). International implementation scheme, United Nations decade of education for sustainable development. 2005-2014. UNESCO. www.unesco.org/education

Veiga Ávila, L., Rossato Facco, A. L., Bento, M. H. D. S., Arigony, M. M., Obregon, S. L., & Trevisan, M. (2018). Sustainability and education for sustainability: An analysis of publications from the last decade. *Environmental Quality Management, 27*(3), 107-118.

Velazquez, L., Munguia, N., & Sanchez, M. (2005). Deterring sustainability in higher education institutions: An appraisal of the factors which influence sustainability in higher education institutions. I*nternational Journal of Sustainability in Higher Education, 6*(4), 383-391.

Velazquez, L., Munguia, N., Platt, A., & Taddei, J. (2006). Sustainable university: what can be the matter?. *Journal of cleaner production, 14*(9-11), 810-819.

Vilmala, B. K., Karniawati, I., Suhandi, A., Permanasari, A., & Khumalo, M. (2022). A Literature Review of Education for Sustainable Development (ESD) in Science Learning: What, Why, and How. *Journal of Natural Science and Integration, 5*(1), 35-44.

Wang, W. (2015). An exploration of patterns in the practice of education for sustainable development in China: experience and reflection. *Open Journal of Social Sciences, 3*(05), 64.

Wals, A. E., & Lenglet, F. (2016). Sustainability citizens: Collaborative and disruptive social learning. In Sustainability citizenship in cities (pp. 52-66). Routledge.

Weiss, M., Barth, M., & von Wehrden, H. (2021). The patterns of curriculum change processes that embed sustainability in higher education institutions. *Sustainability Science, 16*(5), 1579-1593.

Westermark, Å., & Jansund, B. (2019). Learning experiences from a time-geographic approach—Commodity chains, globalization, everyday life, and sustainability in context. *Journal of Geography in Higher Education, 43*(4), 486-504.

Widodo, A., Kaniawati, I., & Fujii, H. (2023). The Development and Validation of an Instrument for Assessing Science Teacher Competency to Teach ESD. *Sustainability, 15*(4), 3276.

Wiek, A., Withycombe, L., & Redman, C. L. (2011). Key competencies in sustainability: a reference framework for academic program development. *Sustainability science, 6*, 203-218.

Wright, T., & Pullen, S. (2007). Examining the literature: A bibliometric study of ESD journal articles in the Education Resources Information Center Database. *Journal of Education for Sustainable Development, 1*(1), 77-90.

Yarime, M., & Tanaka, Y. (2012). The issues and methodologies in sustainability assessment tools for higher education institutions: a review of recent trends and future challenges. *Journal of Education for Sustainable development, 6*(1), 63-77.

iejee

International Electronic Journal of Elementary Education
September 2023, Volume 16, Issue 1, 89-98

A Test Development Study on Spatial Visualization for Second-Grade Primary School Students

Sevgi Demirel[a], Hatice Cetin[b,*]

Received : 2 June 2023
Revised : 11 September 2023
Accepted : 29 September 2023
DOI : 10.26822/iejee.2023.316

[a] Sevgi Demirel, Ministry of National Education, Türkiye.
E-mail: sevgidemirel129@gmail.com
ORCID: https://orcid.org/0000-0002-5853-1344

[b,*] **Corresponding Author:** Hatice Cetin, Necmettin Erbakan University, Konya, Türkiye.
E-mail: haticebts@gmail.com
ORCID: https://orcid.org/0000-0003-0686-8049

Abstract

This study aims to develop a new spatial visualization test (SVT) for second grade primary school students. The study employed the survey design, and the test was developed in accordance with the test development steps. According to the findings obtained as a result of the pilot study, the items were generally high difficulty levels, and they were very good items. Exploratory and confirmatory factor analysis findings confirmed the three sub-dimensions (mental integration, mental rotation, and paper folding) measuring spatial visualization skills. SVT, which was reduced to 10 items, was administered to 396 students; the KR-20 internal consistency coefficient was 0.63, the average difficulty index was 0.35 and the average discrimination index was 0.48. As a result of the study, a valid and reliable test on spatial visualization for second grade primary school students was obtained and it was concluded that students' spatial visualization was at a low level.

Keywords:

Mental Rotation, Paper Folding, Spatial Visualization, Test development, Factor, Tetrachoric factor analysis.

Introduction

Frequently used in daily life, spatial skills are a set of skills based on envisioning and mental visualization. This ability is very critical for solving real-world problems and is also employed in identifying the problem, drawing an organized path, and determining the solution steps (Turğut & Yılmaz, 2012). Hendroanto et al. (2015) defined the concept of spatial ability as the ability to understand, manipulate, reorganize, and interpret visually and summarized all the definitions of spatial ability and briefly stated that it is a mental ability. The National Council of Teachers of Mathematics (NCTM, 2000) noted that spatial ability is a fundamental skill for students and emphasized the importance of this concept.

The availability of several definitions of spatial ability has brought along many components. An analysis of published research suggests that researchers used different components in their studies. Lohman (1996) mentioned the existence of two major components in his study. These components are spatial relations and spatial visualization. A different classification was made by McGee (1979) and

Tartre (1990). McGee (1979) stated that the concept of spatial ability consists of a combination of spatial visualization as the most generally accepted component and spatial orientation in addition to it, as in many previous studies. Tartre (1990), based on McGee's study, referred to the spatial ability concept consisting as a combination of visualization and orientation factors. Similarly, Lord (1985) mentioned the existence of spatial visualization and spatial orientation components for the concept of spatial ability. Linn and Petersen (1985) also mentioned four different components and one of the components was spatial visualization. As it is understood from all these studies on spatial ability, the instruments that measure spatial ability skills appear with the name or component of spatial visualization. Accordingly, the concept of spatial ability and the concept of spatial visualization have been used interchangeably.

McGee (1979) defined the concept of spatial visualization as one's manipulation of a picture presented to an individual in his/her mind, and rotating, folding, or reversing it. Indeed, spatial visualization is a complex concept that includes spatial orientation and mental rotation (Linn & Petersen, 1985). A review of literature presents many spatial visualization tests. The most generally accepted test among these is the "MGMP Spatial Visualization Test (Middle Grades Mathematics Project: Spatial Visualization)" developed by Winter et al. (1989) and adapted into Turkish by Turğut (2007). Other spatial visualization tests in the literature include Purdue Spatial Visualization Test (PSVT) (Guay, 1977), Embedded Figures Test (Witkin et al., 1977), Surface Development Test (Thurstone & Thurstone, 1949), Daily Occupational Test (Eliot & Smith, 1983), and Cube Formation Test (Alias et al., 2002). The questions used to measure spatial visualization skills in Dokumacı Sütçü and Oral's (2019) study and the spatial visualization questions in Hawes et al.'s (2017) study were taken as reference and different types of questions measuring spatial visualization skills were used in the current study.

Different question patterns were used to measure spatial visualization skills in the relevant literature. Mental integration and mental decomposition questions (Hawes et al., 2017; Dokumacı Sütçü & Oral, 2019); paper folding questions (Ekstrom, 1976; French et al., 1963; Kyllonen et al., 1984; Linn & Petersen, 1985; Thurstone & Thurstone, 1949) and mental rotation questions (Guay, 1977; Quaiser-Pohl, 2003; Shepard & Metzler, 1971) are the questions used to measure spatial visualization skill. In the spatial visualization skill test developed for this study, mental integration, paper folding, and mental rotation questions were used to measure this skill.

Reviewing the literature on mental integration questions, we observed that the most appropriate study for primary school level was conducted by Hawes et al. (2017) and Dokumacı Sütçü and Oral (2019). There is a wider network of publications in the literature on paper folding questions. The paper folding test, developed by Ekstrom (1976) and adapted into Turkish by Delialioğlu (1996), has been accepted as one of the important tests in measuring this skill. On the other hand, Linn, and Petersen (1985), Kyllonen et al. (1984) and French et al. (1963) also tried to measure spatial visualization skills with paper folding questions. Paper folding questions involve folding the paper once in a certain direction, punching holes in it and determining its final appearance. The most notable studies in the literature on mental rotation questions include Quaiser-Pohl (2003), Guay (1977), Shepard and Metzler (1971) and Peters et al. (1995). In these studies, the test questions were designed as questions requiring the ability to rotate shapes or objects in two- or three-dimensional space in the mind and then to rotate/animate the resulting image in the mind. The above-mentioned tests are mostly old and have been used in the field of technology and engineering, and do not address students at primary school level. The lack of adequate spatial visualization tests in the literature addressing the levels of primary school students has revealed the need to develop a new test. To this end, this study was conducted to develop an up-to-date and original test suitable for the level of primary school students.

In the proposed test, mental integration questions were prepared making use of the questions in the studies of Hawes et al. (2017) and Dokumacı Sütçü and Oral (2019); mental rotation questions were prepared making use of the questions in the studies of Hawes et al. (2017), Lowrie et al. (2019) and Quaiser-Pohl (2003); and paper folding questions were prepared making use of the questions in the study of Linn and Petersen (1985). Thus, spatial visualization questions were prepared for the sample determined within the scope of the research in accordance with the second-grade level through a comprehensive literature review. This Spatial Visualization Test (SVT) aims to diagnose and determine the spatial visualization skills of second-grade primary school students. Accordingly, this research study sought to answer the following research questions:

1- What are the validity and reliability findings of "Spatial Visualization Test for Second-Grade Primary School Students" developed in the study?

2- What is second grade students' spatial visualization level?

3- Is there a significant difference between second grade students' spatial visualization sub-scores?

Method

Research Design

This study, which aims to develop a test to measure the spatial visualization skills of second-grade students, employed the cross-sectional survey model. A cross-sectional study occurs at one point in time and provides a picture of what the researcher wants to study (Allen, 2017). The main purpose of survey studies is to produce quantitative data and various statistics about situations and events over large groups (Creswell, 2009). In this study, students' spatial visualization skills were measured at one point in time with the SVT, developed by the researchers.

Research Population and Sample

The population of the study consisted of second-grade primary school students studying in the 2021-2022 academic year. The sample of the study consisted of 396 second-grade primary school students studying in different public schools in the provinces/districts of Türkiye at each academic achievement level. The 396 students constituting the sample were selected from the accessible population through the simple random sampling method. In this method, the selection process for the sample from the population is carried out in accordance with the principle of randomness, that is, the selection status of the units that can be selected for sampling is independent of each other (Büyüköztürk et al., 2014). Participants were selected from provinces in the eastern, central, and western regions of the country. In addition, while forming the sample, 10% of the accessible population, 10 times the number of items in the test, or more than 10 times the number of test items were reached based on the sample number resulting from the G-power analysis (Pallant, 2020).

Ethical Approval

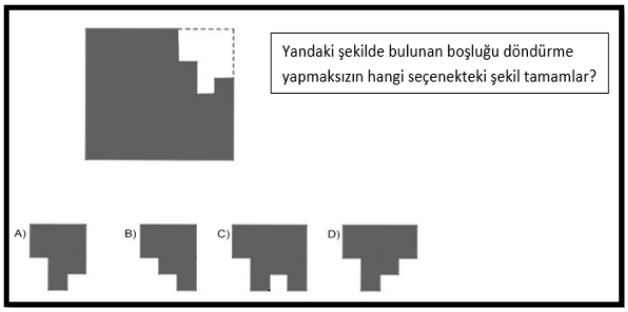
The ethical approval was obtained from the Social Sciences and Humanities Scientific Research Ethics Committee of Necmettin Erbakan University, dated 16.09.2021 with the decision number 2021/448.

Instrument Development Process

The SVT, which was used as a data collection tool, was developed in accordance with the steps of test development including the steps of determining the purpose of the test, determining the scope of the test, determining the question type, determining the number and duration, determining the validity of the test, piloting the test, reliability, and item analysis (Özçelik, 2013).

The aim of the test is to determine the spatial visualization skills of second-grade students. To form the scope of the developed test, the sub-dimensions of the test were determined in the light of the relevant literature and the learning outcomes in the primary school mathematics curriculum. Four multiple-choice questions were prepared for each sub-dimension (mental integration, mental rotation, and paper folding). The developed questions were categorized as easy, medium, and difficult according to their difficulty levels. Mental integration questions, including two easy, one difficult and one medium level questions, were developed by taking the questions in Hawes et al. (2017) and Sütçü and Oral's (2019) studies as examples. Mental rotation questions were developed as two difficult, one easy, and one intermediate level questions based on Hawes et al. (2017), Lowrie et al. (2019), and Quaiser-Pohl (2003). The paper folding questions were prepared making use of Linn and Petersen's (1985) study with two difficult, one easy and one intermediate level questions. Sample items belonging to the sub-dimensions are presented below Figure 1, Figure 2 and Figure 3:

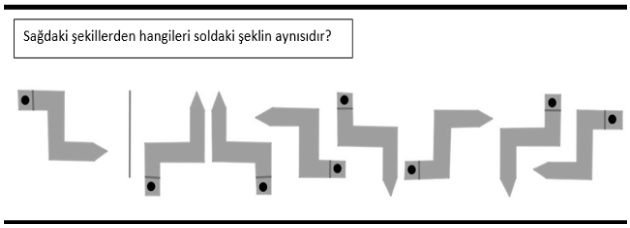
Figure 1.
Sample item for mental integration



(Which option completes the gap in the adjacent figure without making a rotation?)

The mental integration questions consisted of questions requiring the ability to complete incomplete shapes and to combine small shapes into a large shape, as shown in Figure 1.

Figure 2.
Sample item for mental rotation

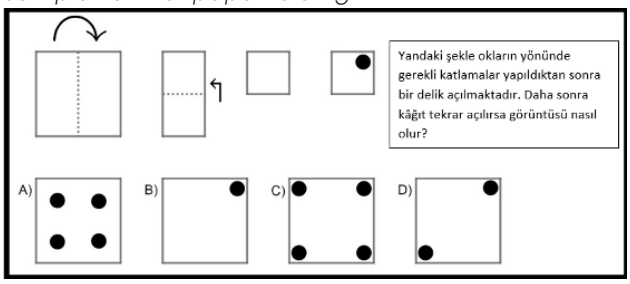


(Which of the figures on the right is the same as the figure on the left?)

Mental rotation questions consisted of questions requiring the ability to determine the new state of the object by performing mental rotation in accordance with the instructions as shown in Figure 2.

Figure 3.

Sample item for paper folding



(After making the necessary folds in the direction of the arrows in the adjacent figure, there will be a hole in the paper. If the paper is then opened again, what will it look like?)

Folding lines and folding directions are shown in the paper folding questions as shown in Figure 3. All paper folding questions in the test consisted of questions that required folding skills in accordance with the instructions.

Data Collection and Analysis

Each participant's test was numbered and recorded using the Microsoft Office Excel program. Each question answered correctly in the test was recorded as 1 (one) point and each question answered incorrectly or left blank was recorded as 0 (zero) point. The highest score to be obtained from the test was 12 and the lowest score to be obtained was 0. Exploratory and confirmatory factor analyses were conducted to ensure the construct validity of the test. The tetrachoric factor analysis was performed with the "Factor" program developed by Lorenzo-Seva and Ferrando (2006). This program provides both Exploratory Factor Analysis (EFA) and Confirmatory Factor Analysis (CFA) results. The descriptive statistics of the students' spatial visualization total scores and the scores of the spatial visualization sub-dimensions were analyzed with the SPSS 22.0 package program. All these analyses conducted as a result of the application of this developed test were carried out to obtain better results and contribute to the literature by giving accurate and consistent results.

Validity and Reliability

The developed test was first piloted with 296 students. As a result of the pilot application, the test was finalized, and applied with the participants. For the content validity, a specification table for the objectives in the curriculum was prepared before the application, a question pool was created in the light of the relevant literature, and then expert opinion was obtained. In line with the opinions of two mathematics education experts and a primary school teacher, the coefficient of agreement was calculated as 0.90. Criterion and construct validity of the study were also checked. Difficulty and discrimination indices were also examined for the validity study of the test. The item difficulty index (p) takes a value between "0" and "1" and when this value is close to 1, it means that the questions are easy and when it is close to 0, it means that the questions are difficult. It is ideal for the questions to be of medium difficulty. The difficulty index values of the items in the measurement tool should be between 0.3 and 0.7 (Büyüköztürk et al., 2014).

The item discrimination index (r) takes a value between "-1" and "+1", and a value close to 1 indicates a higher level of discrimination. A negative discrimination index means that the question is reverse discriminative, that is, the students in the lower group are more successful than the students in the upper group, which indicates that the question does not work. The discrimination index should be 0.30 and above, questions below 0.30 should be corrected, questions with a discrimination level below 0.20 and negative questions should be discarded from the measurement tool (Büyüköztürk et al., 2014). The discrimination index is calculated by subtracting the number of participants from the upper group who answered the question correctly from the number of participants from the lower group who answered the question correctly and dividing by the number of participants in the upper or lower group.

Exploratory and confirmatory factor analyses were conducted to measure the construct validity of the test. Factor analysis is of two types: exploratory factor analysis and confirmatory factor analysis. Scales have equal intervals and achievement tests have an ordinal structure. Therefore, since achievement tests are ordinal scales coded by giving "1" point to the correct answer and "0" point to the wrong answer according to the classical test theory, it is appropriate to perform tetrachoric factor analysis in determining the factor structure, unlike the scales. Within the scope of EFA, KMO value and sig. value of Barlett's test are given. The KMO value obtained because of factor analysis is an indicator of both sampling adequacy and normal distribution. The KMO value should be 0.60 and above (Seçer, 2017). The fact that the sig. value in the KMO table is less than 0.05 means that the data set is suitable for conducting factor analysis and the data are significant. When these two conditions are met, factor analysis data can be interpreted. The Eigenvalues of the factors are expected to be "1" and above, and all factors are expected to explain at least 40% of the total variance (Seçer, 2017). Within the scope of CFA, Chi-squared/ Degrees of Freedom and other fit index values are calculated. For the factor structure obtained as a result of EFA to be confirmed in CFA, the "Chi-squared/ Degree of Freedom" value is expected to be less than 5 (Seçer, 2017). The validity and reliability analysis findings of the test are given in the Findings section.

Findings

The findings regarding the research question, "What are the validity and reliability findings of "Spatial Visualization Test for Second-Grade Primary School Students" developed in the study?", are provided below.

Findings regarding the Validity of the Spatial Visualization Test

Exploratory factor analysis and confirmatory factor analysis were used for construct validity in the study. Table 1 shows the exploratory factor analysis findings of the test.

Table 1.
Exploratory Factor Analysis Results of the Spatial Visualization Test

Item	Factor 1 (M. I.)	Factor 2 (M. R.)	Factor 3 (P. F.)
1	0.312		
4	-		
8	0.529		
11	0.353		
2		0.525	
5		0.729	
7		0.812	
10		0.992	
3			0.406
6			0.598
9			-
12			0.855
			Total Variance: % 61.549

The exploratory factor analysis results presented in Table 1 demonstrated that the 61% variance in the test was explained with three sub-dimensions. The factor loadings of the items related to these factors ranged between 0.312 and 0.529 in the mental integration sub-dimension, between 0.525 and 0.992 in the mental rotation sub-dimension, and between 0.406 and 0.855 in the paper folding sub-dimension. The item factor load should be at least .30 (Seçer, 2017, p. 166). Since items 4 and 9 were not explained under any factor, we decided to remove them from the test. The findings of the confirmatory factor analysis of the developed test are given in Table 2.

Table 2.
Confirmatory Factor Analysis Results of the Spatial Visualization Test

Fit Indices	Observed Values	Result
χ^2	87.421	-
$\chi 2$/ df	3.49	Excellent Fit
RMSEA	0.064	Excellent Fit
GFI	0.972	Excellent Fit
AGFI	0.938	Excellent Fit
NNFI	0.956	Excellent Fit
CFI	0.980	Excellent Fit

As presented in Table 2, 0.90 was regarded as the acceptable cut-off value and 0.95 was regarded as the excellent cut-off value for the Chi-square Goodness, GFI (Goodness of Fit Index), RMSEA (Root Mean Square Error of Approximation), CFI (Comparative Fit Index) and AGFI (Adjusted Goodness of Fit Index) fit indices for confirmatory factor analysis (Meydan & Şeşen, 2011). For RMSEA, a value less than 0.08 was accepted as an acceptable fit value. According to the confirmatory factor analysis findings, the fit indices were found to be significant (x^2 = 87.421, df = 25, p > 0.05; x^2/df=3.49). The fit index values were RMSEA = 0.064, NFI = 0.956, CFI = 0.980, AGFI=0.938, GFI = 0.972. Accordingly, the confirmatory factor analysis revealed that the three-dimensional structure gave a good fit. It can be said that the test was developed as a reliable and valid measurement tool with a three-dimensional structure including mental integration, mental rotation and paper folding.

Findings regarding the Item Analysis of Spatial Visualization Test

The values obtained as a result of item analysis to ensure the construct validity of the questions in the developed test are given in Table 3.

Table 3.
Findings Related to Item Analysis of the Spatial Visualization Pilot Test

Item	pj	Difficulty Level	rj	Discrimination Assessment
1	0,922	Easy	0,299	Needs Revision
2	0,543	Medium	0,486	Very good
5	0,203	Difficult	0,569	Very good
4	0,414	Medium	0,378	Good
5	0,203	Difficult	0,295	Needs Revision
6	0,424	Medium	0,502	Very good
7	0,063	Difficult	0,454	Very good
8	0,6	Medium	0,554	Very good
9	0,431	Medium	0,414	Very good
10	0,122	Difficult	0,434	Very good
11	0,564	Medium	0,437	Very good
12	0,354	Difficult	0,527	Very good

As Table 3 demonstrates, most of the test items had a medium-level difficulty. In addition, items 1 and 5 were found to be items whose discrimination values needed to be revised according to the criteria. This pilot test with 296 students revealed that the items were generally at medium and difficult levels and very good items in terms of discrimination. The KR-20 internal consistency coefficient of the test was calculated as 0.57. The average difficulty index of the

test was 0.41 and the average discrimination index of the test was 0.44.

As a result of the pilot application, two items that needed to be corrected were revised (1st and 5th items) and two items were removed from the test according to the factor analysis findings (4th and 9th items) and the finalized 10-item SVT was applied to 396 students. As a result of this application, the construct validity item analysis findings are presented in Table 4.

Table 4.
Item Analysis Results regarding the Final Form of the Spatial Visualization Test

Item number	Pj	Difficulty Level	rj	Discrimination Assessment
1	0,919	Easy	0,260	Needs Revision
2	0,497	Medium	0,448	Very good
3	0,351	Difficult	0,596	Very good
5	0,166	Difficult	0,299	Good
6	0,414	Medium	0,533	Very good
7	0,050	Difficult	0,369	Very good
8	0,542	Medium	0,583	Very good
10	0,118	Difficult	0,434	Very good
11	0,550	Medium	0,451	Very good
12	0,318	Difficult	0,562	Very good

The findings of the analysis of the items presented in Table 4 revealed that the items were generally medium and difficult in terms of difficulty index and very good in terms of discrimination. The KR-20 internal consistency coefficient of the final test was calculated as 0.63. The item discrimination index of the first item was found to be acceptable (very close to 0.30). The average difficulty index of the final test was 0.35 and the average discrimination index was 0.48. The reliability values of the pilot test and the final test are presented in Table 5.

Table 5.
Reliability Analyses of Pilot Test and Final Test

	Item	N	Average Difficulty	Average Discrimination	KR-20
Pilot	12	297	0,41	0,44	0,57
Final Test	10	396	0,35	0,48	0,63

Table 5 shows that the reliability values increased from 0.57 to 0.63 after the two items were removed from the test. As a result of all these analyses and measurements, we can argue that the SVT for second grade primary school students developed by the researchers is a reliable and valid measurement tool with a three-factor structure including mental integration, mental rotation, and paper folding.

The findings related to the second research question, "What is second grade students' spatial visualization level?", are presented below.

Descriptive Statistics regarding Spatial Visualization Test

The findings of the analyses conducted to determine the level of spatial visualization of second grade students are presented in Table 6.

Table 6.
Descriptive Statistics regarding Spatial Visualization Test

	N	Min	Max	Mean	*SD*
Mental Integration	396	0,00	3,00	2,012	0,843
Mental Rotation		0,00	4,00	0,833	0,843
Paper Folding		0,00	3,00	1,083	1,041
Total		0,00	10,00	3,929	2,005

Table 6 shows the participant students' scores for the sub-dimensions of the SVT. Accordingly, the mean total score of the second-grade students for spatial visualization ($\bar{X}$= 3,929; sd=2,005) is low. The mean scores in the mental integration sub-dimension ($\bar{X}$ = 2.01; *sd* = 0.84), mental rotation sub-dimension ($\bar{X}$= 0.83; sd = 0.84), and paper folding sub-dimension ($\bar{X}$= 1.08; sd = 1.04) are below the average. According to the mean spatial visualization scores of the second-grade students, we can argue that the students showed the best performance in the questions belonging to the mental integration sub-dimension and the lowest performance in the questions belonging to the mental rotation sub-dimension.

The findings related to the third research question, "Is there a significant difference between second grade students' spatial visualization sub-scores?", are presented below.

To answer this research question, the results of the repeated measures ANOVA test conducted to determine whether there is a statistically significant difference between the mean scores of the second-grade students in the sub-dimensions of the SVT are presented in Table 7.

A statistically significant difference was found between the second-grade students' mean scores in the sub-dimensions of the spatial visualization test, as presented in Table 7 ([F395,2 = 262,124], p < .05). In terms of the effect size, this significant difference was of medium size (η^2 = 0.399) (Cohen, 1988). As a result of the comparisons made with Bonferonni test, a significant difference was found between mental integration and mental rotation, mental integration and paper folding, mental rotation and paper folding scores (p = .00). Looking at the means of the sub-dimensions, the significant difference between mental integration

Table 7.
Results regarding Spatial Visualization Test

Variables	Variance Source	Sum of Squares	*Sd*	Mean Squares	*F*	*p*	Significant Difference	η^2
Spatial visualization	Between Subjects	529,340	395	1,340				
	Measurement	305,820	2	152,910	262,124	0,00	1-2 1-3 2-3	0,399
	Error	460,847	790	0,583				
	Total	1,296,007	1.187					

p<0.05 significant
1: mental integration 2: mental rotation 3: paper folding

and mental rotation was in favor of mental integration scores ($\bar{X}$ = 2.012; *sd* = 0.843). The significant difference between mental integration and paper folding was also in favor of mental integration scores ($\bar{X}$ = 2,012; *sd* = 0,843). The significant difference between mental rotation and paper folding scores was in favor of paper folding scores ($\bar{X}$ = 1,083; *sd* = 1,041).

Conclusion and Discussion

This study reports the development of a spatial visualization test for second-grade primary school students, and their spatial visualization levels were also investigated. An analysis of the related literature revealed that the tests that require high-level thinking skills such as spatial visualization are simplified and applied to primary school and preschool level students (Dokumacı Sütçü & Oral, 2019; Quaiser-Pohl, 2003). Accordingly, the questions used in the spatial visualization tests in the literature were transformed into a new test to address second-grade students in the current study. The reliability of the Purdue Spatial Visualization Test developed by Guay (1977) was 0.80; the reliability of the Spatial Visualization Test developed by Alias et al. (2002) was 0.55; the reliability of the Spatial Visualization Test in Two-Dimensional Geometry developed by Olkun and Altun (2003) was 0.77; and the reliability of the MGMP Spatial Visualization Test adapted into Turkish by Turğut (2007) was 0.81. The reliability of the Spatial Visualization Test developed in this study was calculated as 0.63. This value suggests that the reliability is quite acceptable when evaluated in terms of tests coded as true-false and the number of items in the test (Seçer, 2017).

The results demonstrated that the sub-dimensions of the SVT developed within the scope of the current study are like the sub-dimensions of Linn and Petersen (1985), Gorska and Leopold (1998) and Burton and Fogarty's (2003) tests. The literature accommodates many spatial visualization definitions (Hauptman, 2010; Hendroanto et al., 2015; Linn & Petersen, 1985; Lohman, 1996), spatial visualization tests (Purdue SVT, MGMP, Daily Occupational Test, Embedded Figures Test), spatial visualization components (Burton & Fogarty, 2003; Caroll, 1993; Gorska & Leopold, 1998; Guilford et al., 1952 Kimura, 1999; Lohman, 1996; Linn & Petersen, 1985; Lord, 1985; McGee, 1979; Tartre, 1990; Thurstone & Thurstone, 1949), indicating that this topic is significant and has attracted the attention of many researchers. However, the very broad perspectives in all these studies have prevented the drawing of clear boundaries for research. It is thought that conducting studies within clear boundaries will be more beneficial for other researchers. The scope of the test in this study was composed of mental rotation, mental integration and paper folding factors, and was confirmed by exploratory and confirmatory factor analyses.

Significant differences were found between second-grade students' spatial visualization scores among the three sub-dimensions. The second-grade students scored the highest in the questions in the mental integration sub-dimension and the lowest in the questions belonging to the mental rotation sub-dimension. In line with this result, mental rotation was determined as the skill that students had the most difficulty within the studies of Yılmaz and Yenilmez (2019) and Turgut and Yılmaz (2012). In Sezen Yüksel's (2013) study, participants' mental rotation scores were found to be higher than other sub-dimensions. There was a significant difference between mental integration skills and paper folding skills in favor of the mental integration sub-dimension. On the other hand, there was a significant difference between paper folding skills and mental rotation skills in favor of paper folding scores. Among these three sub-dimensions, the sub-dimension in which the participants were most accurate was determined as the mental integration sub-dimension. Lowrie et al. (2019) put forward an opinion that explains this situation in their study; mental rotation is a more complicated skill that requires a higher level of spatial visualization. In addition, Lowrie et al. (2019) argue that students should have skills such as paper folding to answer mental rotation questions. Accordingly, we can argue that paper folding is a more primal and prerequisite skill than mental rotation. The result of another study, which contradicts this result, is that students' mental rotation skill averages were higher than the other sub-dimensions of the study (Sezen Yüksel, 2013). According to Sezen Yüksel (2013), mental rotation skill is a prerequisite for spatial

visualization and mental cutting sub-dimension. It was also found that mental rotation ability explained 90% of spatial ability, while spatial visualization ability explained 50% of spatial ability. According to Çakmak et al. (2014), 10% of spatial visualization can be explained by origami-based education enriched with paper folding activities.

Recommendations and Limitations

Future research may investigate the parameters according to which the spatial visualization sub-dimension scores differ. The relationship between this test and different mathematical skill variables can also be investigated. In this study, a factor related to three-dimensional elements was not identified because three-dimensional thinking is not fully developed in second grade students. We suggest that future research may develop the test further by adding a three-dimensional thinking factor at different grade levels. The study's limitations include the fact that it was conducted in a few regions in Turkey and focused solely on the primary school level. For practitioners, we recommend using this test in diagnosing second-grade students' spatial visualization, determining their spatial visualization levels and in the developing their spatial visualization skills. In addition, by practicing with items parallel to the items in the spatial visualization test, students' spatial-abstract thinking can be improved. A comparative analysis of the study, which utilizes tests working with different dimensions, and a discussion of the sub-dimensions of the spatial visualization test will contribute significantly to the related field.

Declaration of the Author(s)

Declaration of researchers' contribution rate: The authors contributed equally to the study.

Ethical Committee Decision: The ethical approval was obtained from the Social Sciences and Humanities Scientific Research Ethics Committee of Necmettin Erbakan University, dated 16.09.2021 with the decision number 2021/448.

Conflict of interest: The authors declare no conflict of interest.

Support and thanks: The authors would like to thank the primary school students and their esteemed teachers who participated in the study.

References

Alias, M., Black,T.R., Gray, D.E. (2002), Effect of Instructions on Spatial Visualisation Ability in Civil Engineering Students, *International Education Journal*, 3 (1), 1-12.

Allen, M. (Ed.). (2017). *The SAGE encyclopedia of communication research methods*. SAGE publications.

Burton, L., & Fogarty, G. (2003), The Factor Structure of Visual Imagery And Spatial Abilities, *Intelligence*, 31, (3), 289-318. https://doi.org/10.1016/S0160-2896(02)00139-3

Büyüköztürk, Ş., Kılıç-Çakmak, E., Akgün, Ö., Karadeniz, Ş., & Demirel, F. (2014), *Bilimsel Araştırma Yöntemleri*, Pegem Akademi Yayıncılık.

Carroll, J. B. (1993). *Human cognitive abilities: A survey of factor-analytic studies* (No.1), Cambridge University Press

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Second Edition. Hillsdale, NJ: Lawrence Erlbaum Associates, Publishers

Creswell, J. W. (2009). "Mapping the field of mixed methods research", *Journal of mixed methods research,* 3(2), 95-108

Çakmak, S., M. Işıksal ve Y. Koç (2014)." Investigating effect of origami-based instruction on elementary students spatial skills and perceptions", *Journal of Educational Research*, 107(1), 59–68. https://doi.org/10.1080/00220671.2012.753861

Delialioğlu, Ö. (1996), *Contribution of Students' Logical Thinking Ability on Achievement in Secondary Physics*, Master Degree Thesis, ODTÜ Fen Bilimleri Enstitüsü, Ankara.

Dokumacı Sütçü, N., & Oral, B. (2019), Uzamsal Görselleştirme Testinin Geliştirilmesi: Geçerlik ve Güvenirlik Çalışmaları, *Kastamonu Eğitim Dergisi*, 27(3), 1179–1195. https://doi.org/10.24106/kefdergi.2826

Ekstrom, R. B. (1976), Kit of factor-referenced cognitive tests, *Educational Testing Service*.

Eliot, J.,& Smith, I.M., (1983), *An International Directory of Spatial Tests*, Windsor, The NFER-Nelson Publishing Company Ltd., United Kingdom.

French, J. W., Ekstrom, R. B., & Price, L. A. (1963). *Manual for kit of reference tests for cognitive factors* (revised 1963). Educational Testing Service Princeton NJ.

Gorska, R., & Leopold, C. (1998), Gender differences in visualization skills – An international perspective, *Engineering Design Graphics Journal*, 62(3), 9–18.

Guay, R. B. (1977), Factors Affecting the Development of Two Spatial Abilities Basic to Technical Drawing, *Journal of Industrial Teacher Education*, 14(3), 38-43.

Guilford, J. P., Fruchter, B. and Zimmerman, W. S. (1952), Factor analysis of the Army Air Forces Shepard Field battety of experimental aptitude tests, *Psychometrika*, 17: 45–68.

Hauptman, H. (2010). "Enhancement of spatial thinking with Virtual Spaces 1.0.", *Computers & Education*, 54(1), 123-135. https://doi.org/10.1016/j.compedu.2009.07.013

Hawes, Z., Moss, J., Caswell, B., Naqvi, S., & MacKinnon, S. (2017), Enhancing Children's Spatial and Numerical Skills through a Dynamic Spatial Approach to Early Geometry Instruction: Effects of a 32-Week Intervention, *Cognition and Instruction*, 35(3), 236– 264 https://doi.org/10.1080/07370008.2017.1323902

Hendroanto, A., Budayasa, I. K., Abadi, A., Galen, F. V., & Van Eerde, H. A. A. (2015). Supporting Students? Spatial Ability In Understanding Three-Dimensional Representations. In: The Third South East Asia Design/Development Research International Conference, 18 Apr - 19 Apr 2015, Palembang.repository.unsri.ac.id

Kyllonen, P.C., Lohman, D.F., & Snow,R.E., (1984), Effects of Aptitudes, Strategy Training, and Task Facets on Spatial Task Performance, *Journal of Educational Psychology*, 76 (1), 130-145. https://doi.org/10.1037/0022-0663.76.1.130

Linn, M. C., & Petersen, A. C. (1985), Emergence and characterization of sex differences in spatial ability: A meta-analysis, *Child development*, 1479-1498. https://doi.org/10.2307/1130467

Lohman, D. F. (1996), *Spatial Ability and G. Human Abilities: Their Nature and Assessment*, DP Tapsfield, 97–116.

Lord T.R., (1985), Enhancing The Visuo-Spatial Aptitude of Students, *Journal of Resarch in Science Teaching*, 22:395-405. https://doi.org/10.1002/tea.3660220503

Lorenzo-Seva, U., & Ferrando, P. J. (2006), FACTOR: A computer program tofit the exploratory factor analysis model, *Behavior research methods*, 38(1), 88-91.https://doi.org/10.3758/BF03192753

Lowrie, T., Logan, T., & Hegarty, M. (2019), The Influence of Spatial Visualization Training on Students' Spatial Reasoning and Mathematics Performance, *Journal of Cognition and Development*, 20(5), 729–751. https://doi.org/10.1080/15248372.2019.1653298

McGee, M. G., (1979), Human Spatial Abilities: Psychometric Studies And Environmental, Genetic, Hormonal, And Neurological Influences, *Psychological Bulletin*, 86 (5), 889-918. https://doi.org/10.1037/0033-2909.86.5.889

Meydan, C. H., & Şeşen, H. (2011). *Yapısal eşitlik modellemesi* AMOS uygulamaları. Detay Yayıncılık.

NCTM (National Council of Techers of Mathematics) (2000), *Principles and standards for School Mathematics*. Reston. VA: Author

Olkun, S., & Altun, A. (2003), İlköğretim öğrencilerinin bilgisayar deneyimleri ile uzamsal düşünme ve geometri başarıları arasındaki ilişki, *The Turkish Online Journal of Educational Technology*, 2(4), 86-91.

Özçelik, D.A. (2013). *Test Hazırlama Kılavuzu*, Pegem Akademi, Ankara.

Pallant, J. (2020). *SPSS Survival Manual: A Step by Step Guide to Data Analysis Using IBMSPSS*, Routledge.

Peters, M., Laeng, B., Latham, K., Jackson, M., Zaiyouna, R. and Richardson, C. A Redrawn (1995), Vandenberg & Kuse Mental Rotations Test: Different Versions and Factors that affect Performance, *Brain and Cognition*, 28, 39-58. https://doi.org/10.1006/brcg.1995.1032

Quaiser-Pohl, C., (2003), The Mental Cutting Test "Schnitte" and Picture Rotation Test- Two New Measures to Assess Spatial Ability, *International Journal of Testing*, 3 (3), 219- 231. https://doi.org/10.1207/S15327574IJT03032

Seçer, İ. (2017). *SPSS ve LISREL ile Pratik Veri Analizi: Analiz ve Raporlaştırma*, Anı Yayıncılık.

Sezen Yüksel, N. (2013), *Uzamsal Yetenek Bileşenleri ve Uzamsal Yeteneğin Geliştirilmesi Üzerine*, Doktara Tezi, Hacettepe Üniversitesi, Ankara.

Shepard, R.N.,& Metzler, J. (1971), Mental Rotation of Three- Dimensional Objects, *Science New Series*, 171, 701- 703. https://doi.org/10.1126/science.171.3972.701

Tartre, L.A., (1990), Spatial Orientation Skill and Mathematical Problem Solving, *Journal for Research in Mathematics Education*, 21 (3), 216- 229. https://doi.org/10.5951/jresematheduc.21.3.0216

Thurstone, T. G., & Thurstone, L. L. (1949), *Mechanical aptitude II description of group tests*, Chıcago Unıv Il Psychometrıc Lab.

Turğut, M. (2007), *İlköğretim II. kademede öğrencilerin uzamsal yeteneklerinin incelenmesi*, Yayımlanmamış yüksek lisans tezi, Dokuz Eylül Üniversitesi, İzmir.

Turğut, M., & Yılmaz, S. (2012), İlköğretim 7. ve 8. Sınıf Öğrencilerinin Uzamsal Yeteneklerinin İncelenmesi, *Dicle Üniversitesi Ziya Gökalp Eğitim Fakültesi Dergisi*, (19), 69-79.

Winter, J. W., Lappan, G., Fitzgerald, W. ve Shroyer, J., (1989), *Middle Grades Mathematics Project: Spatial Visualization*, NY: Addison-Wesley.

Witkin, H. A., Moore, C. A., Goodenough, D. R., & Cox, P. W. (1977), Field-dependent and field-independent cognitive styles and their educational implications, *Review of educational research*, 47(1), 1-64. https://doi.org/10.3102/00346543047001001

Yılmaz, E.B. & Yenilmez, K. (2019, September 26-28), *Ortaokul 7. ve 8. sınıf öğrencilerinin Uzamsal Görselleştirme Becerilerinin İncelenmesi*, 4th International Symposium of Turkish Computer and Mathematics Education, İzmir – Full Test

iejee

International Electronic Journal of Elementary Education
September 2023, Volume 16, Issue 1, 99-109

The Impact of Self-Esteem on Teacher Leadership: An Experimental Design

Özgür Bolat[a,*]

Received : 8 June 2023
Revised : 12 September 2023
Accepted : 29 September 2023
DOI : 10.26822/iejee.2023.317

[a] **Corresponding Author:** Özgür Bolat, Faculty of education, International Final University, Girne, TRNC.
E-mail: ozgur.bolat@final.edu.tr
ORCID: https://orcid.org/0000-0003-4420-1368

Abstract

This study aims to examine the impact of self-esteem on teacher leadership beliefs in Türkiye through both correlational and experimental analyses since self-esteem is considered to be a crucial antecedent of leadership. The author conducts three separate studies with a sample size of 808. The first study (n = 237) uses correlational analyses (Spearman). The second study (n = 222) adopts a scenario-based, within-participants experimental design where the self-esteem level of teachers is manipulated. A paired t-test was used to analyze the data. The third study (n=349) employs a between-participants experimental design. ANOVA test was conducted to analyze the data. The correlational analysis of the first study found a positive relationship between self-esteem and teacher leadership (rho =.35, p < .001). The second study revealed that participants perceived teachers with high self-esteem to be more likely to exercise leadership, $t_{(221)} = 1.97$, $p < 0.01$. The third study found that teachers who were primed to have high self-esteem were more likely to adopt leadership beliefs, $F_{(2, 348)} = 3.91$, $p = .02$. The result experimentally demonstrated that self-esteem is an important antecedent of teacher leadership. This study contributes to the literature on teacher leadership by providing a causal link between self-esteem and teacher leadership.

Keywords:

Teacher Leadership, Self-Esteem, Experimental Design, School Improvement, School Leadership.

Introduction

In today's ever-changing world, schools face unprecedented demands and challenges. On the one hand, schools are pressured to prepare students for high-stakes testing in many (developing) countries, such as Türkiye. On the other hand, schools are expected to develop students' skills essential for the 21st century, such as critical thinking, creativity, problem solving and more (OECD, 2013). Furthermore, school improvement is no longer about improving student grades, test results or even students' skills for the 21st century. As Ainscow et al. (1998) highlighted years ago, school improvement requires profound change at all levels of the school system. This encompasses changes at pedagogical level, promoting a positive school culture, empowering teachers, transformational leadership from principals, mobilization of parents, and adoption of supportive educational policy

ISSN: 1307-9298

level (Bolat, 2013). Taking these multifaceted demands into account, principals cannot achieve the expected level of school improvement alone (Shen et al., 2020) especially in countries, such as Türkiye where there is a strong hierarchical school system (Gümüş et al., 2021). Therefore, principals looked to teachers for change in classrooms (Margolis and Huggins, 2012). Furthermore, in the era of accountability, teacher leadership has been promoted as a key strategy for enhancing student results (Lovett, 2018; Wenner and Campbell, 2017). The Teacher Leadership Exploratory Consortium (2011) proposed teacher leadership as an effective strategy for school improvement. There is now enough empirical evidence that teacher leadership contributes to the quality of schools (Wang and Ho, 2020) and school improvement (Ingersoll et al., 2018; Shen et al., 2020; Poekert et al., 2016; Wenner and Campbell, 2017). Therefore, there is a global movement towards promoting teacher leadership within schools. In line with these continuous efforts, this study aims to make a valuable contribution by providing experimental evidence that self-esteem could be a crucial antecedent for teacher leadership. By enhancing teachers' self-esteem, school leaders could enhance and promote teacher leadership within educational settings. Similarly, teachers can work on improving their self-esteem to enhance their capacity for effective leadership.

Literature Review

Teacher Leadership

Although teacher leadership is defined differently (Nguyen et al., 2020), Wenner and Campbell's (2017) definition summarizes the essence of most definitions of teacher leadership most concisely: "teachers assuming leadership roles beyond their classrooms while simultaneously fulfilling their classroom responsibilities." These roles could be "positional", such as department head, curriculum designer, instructional specialist, mentor and etc. (see Katzenmeyer and Moller, 2009; Neumerski, 2012) or "non-positional" (Frost, 2019), such as leading professional development activities, participating in decision making, assisting colleagues, observing colleagues in classrooms and etc. (see Wenner and Campell, 2017). In other words, teachers could perform these leadership roles informally or formally (Meirink et al., 2020). Some scholars, such as Silva et al. (2000), argue that when teachers exercise leadership within a formal role, they could be just fulfilling some of principal's management duties, functioning as managers. These formal leadership roles could of course contribute to school improvement. However, such a formal role of teacher leadership, on the contrary, could stifle school improvement, by creating overload, role ambiguity and role conflict (Smylie et al., 2002). Furthermore, formal positions can actually prevent individuals from exercising leadership (Heifetz, 2004) because leadership requires creating tolerable chaos while authority that stems from formal positions requires establishing order and predictability (Heifetz and Linsky, 2017). Teachers in formal positions may find themselves stuck between these two conflicting demands and may prioritize establishing order, which makes it default for them to exercise leadership. Additionally, when teacher leadership is associated with formal roles and positions, teachers without a formal leadership position may be unwilling to exercise leadership (Krieg et al., 2014).

School improvement depends on all teachers' efforts to take responsibility for change and transformation.

In line with this informal view of teacher leadership, the current perspective goes beyond formal roles and encompasses a more informal approach (Carrion and García-Carrión, 2015; Poekert, 2012). In this current perspective, teachers exercise leadership regardless of their positions (Frost, 2019). Leadership is not associated with any title, role or position. Leadership is more about teachers' taking ownership and drive meaningful change within schools (Frost and Durrant, 2002). In a formal leadership position, teachers of course could initiate change but it may not always imply leadership. They could just be solving a problem as part of their job description or leading a project dictated by the principal. Within an informal perspective of teacher leadership, however, leadership extends beyond solving a problem. Leadership is about agency and a sense of autonomy in leading change throughout the school (Durrant and Holden, 2006). In this view, every teacher is regarded as a 'change agent' (Fullan, 2007). The challenge, then, is to explore and understand the antecedents of teacher leadership and foster these antecedents so that every teacher can exercise leadership.

In line with this understanding, researchers are trying to understand the antecedents of teacher leadership (Ding and Thien, 2022). Some of these antecedents are at teacher levels (Schott et al., 2020). Some of these teacher-level antecedents are self-perception (Huang, 2016; Mongillo et al., 2012), experience (Angelle and DeHart, 2011), expertise (Struyve et al., 2018), motivation (Rogers, 2005), skills (Collinson, 2012; Liljenberg, 2016), teacher competence (Supovitz et al., 2010), extraversion (Landis et al., 2022), disposition (Carver, 2016), self-confidence (Pankake and Moller, 2007), and self-efficacy (Sun et al., 2017). An extensive body of research demonstrates that all these antecedents contribute greatly to the enactment of teacher leadership. For instance, teachers with higher levels of self-efficacy are more likely to exercise leadership (Sun et al., 2017). Another important antecedent that has attracted in the literature attention is self-esteem (Hunzicker, 2017; Smulyan, 2016). Therefore, promoting

teachers' self-esteem can be an effective strategy for enabling teachers to exercise leadership.

Self-esteem

The concept of self-esteem is pervasive in contemporary life (Orth and Robins, 2014) and extensively studied in psychology and social sciences (Donnellan et al., 2011; Judge et al., 2002). Despite its prevalence, however, the definitions of self-esteem vary (Heppner and Kernis, 2011). Some scholars define it in terms of affective (feeling) component. Self-esteem is defined as a subjective evaluation of one's worth as an individual (Baumeister, 1999; Donnellan et al., 2011; Rosenberg, 1965; Zeigler-Hill and Showers, 2007). Conversely, some researchers define self-esteem in terms of its cognitive (thought) aspect, such as an individual's overall self-evaluation of their competencies (Korman, 1968). Some scholars define it in terms of both cognitive and affective component, that is in terms of "worthiness" and "competence" (Tafarodi and Swann, 2001). Subsequent studies attempted to clarify this confusion. Some researchers used the term "self-concept" (Campbell, 1990; Lee-Flynn et al., 2011; Huitt, 2004) or "self-efficacy" (Bandura, 1997) to refer to the cognitive aspect of the self. Contemporary understanding highlights that self-esteem is more about feelings rather than thoughts about oneself and not confined to specific abilities or behaviors (Robins et al., 2001). More specifically, self-esteem is concerned about how one "feels" about oneself (Coopersmith, 1967; Sowislo and Orth, 2013). That is, people with high self-esteem like who they are (Pelham and Swann, 1989). This understanding has important implications for teacher leadership. Teachers with little expertise, experience, competence or skills can choose to exercise leadership if they have high self-esteem.

Self-esteem is widely studied because it is believed to affect various significant life outcomes (Swann et al., 2007). A recent meta-analysis found that low self-esteem is related to anxiety (Liu et al., 2022) and depression (Sowislo and Orth, 2013), poor health and criminal behavior (Trzesniewski et al., 2003). Self-esteem affects job satisfaction (Judge and Bono, 2001), relationship satisfaction (Shackelford, 2001), subjective well-being (Diener and Diener, 1995), and positive affect in life (Orth et al., 2012: Robins, 2001). Scholars also study the relationship between self-esteem and leadership since self-esteem is considered a crucial factor in leadership effectiveness (Brockner, 1988). In early years, Beer et al. (1959) and Kipnis and Lane (1962) found that individuals who had high self-esteem emerged as leaders in college groups and among naval officers, respectively. Hill and Ritchie (1977) proposed that self-esteem is predictive of leadership. A meta–analysis by Judge et al. (2002) found a positive correlation between leadership emergence and self–esteem. In a more recent study, Galante and Ward (2017) discovered that female athletes with high self-esteem exhibited more leadership characteristics. Similarly, Matzler et al. (2015) observed a strong positive relationship between transformational leadership and self-esteem in their study of 411 managers and entrepreneurs. In the context of school leadership, Vašašová et al. (2021) found a positive correlation between teacher leadership and self-esteem.

Self-esteem is positively associated with leadership for several reasons. When leaders have high self-esteem, they transmit positivity and enthusiasm more often (Hu et al., 2012). Additionally, individuals with high self-esteem act less defensively when they are subject to negative feedback (Brown, 2010). In contrast, individuals who have low self-esteem get more distressed when they face failure (Holland et al., 2002). Moreover, people with high self-esteem act in a more accepting manner while people with low self-esteem exhibit heightened prejudice (Crocker and Luhtanen, 1990). All these qualities are indicative of effective leadership. Most importantly, individuals need to believe in their capacity to attain power and high self–esteem is a prerequisite of such beliefs (Wojciszke and Struzynska–Kujalowicz, 2007). If they have low self-esteem, they shy away from leadership roles (Brockner, 1988) and avoid challenging tasks (Zeigler-Hill et al., 2011). In sum, an extensive body of research provides evidence that there is a strong relationship between self-esteem and leadership.

Previous research demonstrates a strong relationship between leadership and self-esteem. However, the direction of causality remains unclear since all these studies are correlational. As Baumeister et al. (2003) suggest the causal arrow could point in either direction or Wojciszke and Struzynska–Kujalowicz (2007) suggest there could be reciprocal influences. The current study aims to clarify the nature of this relationship, using an experimental design since experiments are crucial for providing evidence of causality (Antonakis et al., 2010).

Although there is extensive research on teacher leadership, Schott et al. (2020) found in their meta-analysis of 90 studies that none of these studies employed experimental design. Thus, the current study aims to address that gap, using an experimental approach and contribute to the existing literature on teacher leadership. To the best of my knowledge, this study is the first study that investigates the relationship between self-esteem and (teacher) leadership experimentally not only in the field of educational leadership but in psychological /organizational science. School leaders and teachers can use the findings and insights of the study to promote teacher leadership since self-esteem can be improved (Pierce et al.,1989)

Methodology

The current study aims to test the impact of self-esteem on teacher leadership beliefs through three separate studies (N = 808). The first study employed correlational analysis to explore the relationship between self-esteem and teacher leadership beliefs. The second and third studies employed an experimental design to delve deeper into this relationship and establish a causal link.The present study tests three hypotheses:

Hypothesis 1: There is a positive relationship between self-esteem and teacher leadership beliefs (Study 1).

Hypothesis 2: Participants will perceive teachers with a high level of self-esteem as more likely to embrace teacher leadership beliefs, compared to teachers with a low level of self-esteem (Study 2).

Hypothesis 3: Participants who are primed to have high self-esteem will be more likely to adopt teacher leadership beliefs, compared to teachers who are primed to have low self-esteem (Study 3).

Study 1

Participants

The study included a total of 237 participants who were employed as teachers at state schools across various regions in Türkiye. 138 participants (59.2%) were female, while 95 participants (%) were male (Demographic information was not provided by 4 participants). In terms of teaching level, 19 participants (8.1%) were from preschool, 68 participants (29.1%) were from primary school, 73 participants (31.2%) were from middle school, and 74 participants (31.6%) were from high school. The average age of the participants was 45.20 years (range = 28–61, *SD* = 7.4), and they had an average of 20.9 years of teaching experience (range = 1–39, *SD* = 8.1).

Procedure and Measures

Participants completed the abbreviated version of the Rosenberg self-esteem scale (Rosenberg, 1965) consisting of six items (Cronbach's alpha = .84). This shortened version was developed by Bolat and Antalyalı (2023) and demonstrated high levels of validity and reliability. Sample items are: "I feel that I'm a person of worth," "On the whole, I am satisfied with myself," and "I certainly feel useless at times (reverse item)." Additionally, participants also completed a five-item subscale of teacher leadership belief scale (Bolat, 2023). This subscale exhibited strong validity and reliability (Cronbach's alpha = .86). The subscale measures teachers' leadership belief- the extent to which teachers see certain leadership behaviors as part of their professional identity. Sample items are "I engage in pedagogical conversations with my colleagues to contribute to their development" and "I offer advice and suggestions to my colleagues." The responses ranged from 1 (not at all) to 5 (always).

Results and Discussion

Data analysis was performed using Jamovi 2.3.12.0. The Spearman correlation coefficient was used to assess the relationship between the teacher leadership belief scale and the self-esteem scale, since the data did not have a normal distribution (Shapiro-Wilk test, $p < .001$). The analysis indicated that there was a positive correlation between the two scales and the correlation was significant (rho = .35, $p < .001$). This result suggests there is a relationship between these two constructs. Thus, H1 was supported. However, the use of a correlation study fails to establish a causal link between these concepts. It is not possible to tell which concept is the precursor. Therefore, the second study utilized a within-subjects experimental design to establish a causal link between these two constructs.

Study 2

Participants

The experiment included a total of 222 participants who were employed as teachers at state schools across various regions in Türkiye. 185 (88.1%) of the participants were women and 25 (11.9%) were men. The sample size for the study was determined a priori using G*Power (Faul et al., 2009) to achieve a power of 0.80 and an alpha error probability of .05, with a target small-to-medium-sized effect size of .2. Based on a power analysis, a minimum sample size of 199 participants was required. The actual sample size (n = 222) used in the study exceeded this minimum requirement. Among the participants, 35 (16.9%) were employed in kindergarten, 64 (30.9%) in primary school, 44 (21.3%) in middle school, and 64 (30.9%) at the high school The average age of the participants was 41.3 years (range: 23-69, SD = 7.78), and their average work experience was 17 years (range: 1-45, SD = 8.50).

Procedure

Participants were presented with two scenarios describing a teacher with high and low self-esteem, respectively. They were then asked "What is the likelihood that this teacher will..." followed by 6 items (Cronbach's alpha = .94) taken from the micro-level leadership subscale of teacher leadership behavior scale (Bolat and Antalyalı, 2023). The scale assessed the extent to which teachers exercise leadership to influence their colleagues, with sample items including "I engage in pedagogical conversations with my colleagues to contribute to their development" and "I offer advice and suggestions to my colleagues." Participants provided responses on a 5-point scale ranging from 1 (strongly disagree) to 5 (strongly agree).

Results and Discussion

Data analysis was performed using Jamovi 2.3.12.0. A paired t-test was conducted to compare participants' ratings of the two scenarios. The results indicated a significant difference between the first measurement (high self-esteem scenario) and the second measurement (low self-esteem scenario). The mean scores for the first and second measurements were 4.14 (SD = .79) and 2.18 (SD = .82), respectively. This difference was statistically significant (t(221) = 1.97, p < .01). These findings suggest that participants perceived the teacher with high self-esteem as more likely to adopt leadership beliefs compared to the teacher with low self-esteem. This result provides experimental evidence that self-esteem is an important antecedent and determinant of teacher leadership. Based on the results, it could be concluded that teachers who have higher self-esteem are more likely to have leadership beliefs. Thus, H2 was supported. However, it should be noted that a scenario-based experimental studies have certain limitations. Participants may have understood the purpose of the study (social desirability bias) or been affected by the previous scenario (carryover effect), and responded accordingly. Therefore, a between-subjects experimental design is warranted to assess the impact of self-esteem on teacher leadership beliefs. To address that limitation, the next study utilized a between-subjects experimental design.

Study 3

Participants

The study sample comprised 349 teachers employed at state schools across various regions in Türkiye. The sample size for the study was determined a priori using G*Power (Faul et al., 2009) to achieve a power of 0.90 and an alpha error probability of .05, with a target small-to-medium-sized effect size of .2. Based on a power analysis, a minimum sample size of 321 (107 in each group) participants was required. The actual sample size (*n* = 349) used in the study exceeded this minimum requirement.

Data were collected through online platform. Participants were informed about the study, and all participants participated voluntarily in the study. 325 participants (93.1%) were female, while 24 participants (6.9%) were male. In terms of educational levels, 47 participants (13.5%) were from preschool, 94 participants (26.5%) were from primary school, 111 participants (31.8%) were from middle school, and 97 participants (27.8%) were from high school. The average age of the participants was 37.4 years (range = 23–62, *SD* = 6.74), and their average teaching experience was 13.4 years (range = 1–28, *SD* = 7.11).

Procedure

Prior to the experiment, similar to Study 2, participants completed the abbreviated version of the Rosenberg self-esteem scale (Rosenberg, 1965) consisting of six items (Cronbach's alpha = 0.86). Additionally, participants also completed a five-item subscale of teacher leadership behavior scale (Bolat, 2023) (Cronbach's alpha = 0.92).

After completing initial scales, participants were randomly assigned to one of three groups: the high self-esteem group (HSG) (n = 152), the low self-esteem group (LSG) (n = 152), and the control group (n= 115). In the HSG condition, participants were instructed to "generate a list of positive characteristics that they possess," while in the LSG condition, participants were asked to "generate a list of negative characteristics that they possess." These manipulations have previously been used by McGuire and McGuire (1996) in previous research. Participants in the control group were asked to write about their daily routines (adapted from Schmeichel and Vohs, 2009)

After participants wrote their essays, they were instructed to complete one-item self-esteem scale (adapted from Robins et al., 2001). The single-item self-esteem question was "I have high self-esteem." The responses ranged from 1 (definitely does describe me) to 7 (definitely describes me). They were also instructed to respond to two items from the Mood Adjective Checklist (adapted from Raghunathan and Trope, 2002). Two adjectives from the Mood Adjective Checklist were "happy" and "elated." The responses ranged from 0 (definitely does not apply to my feeling at this moment) to 3 (definitely does apply to my feeling at the moment). Finally, participants completed the six-item subscale of the teacher leadership belief scale (Cronbach's alpha = .85) developed by Bolat (2023). The response choices for this subscale ranged from 1 (Strongly disagree) to 5 (Strongly agree). The scores on this subscale served as the dependent variable. After the experiment, participants were thoroughly briefed.

Results

Randomization Check

The Rosenberg self-esteem scale and the teacher leadership behavior scale were utilized for randomization check. An analysis of variance (ANOVA) was conducted to compare the self-esteem scores across the three groups. The results indicated no significant differences in self-esteem scores between the groups ($F(2, 416) = 2.32$, $p = .10$). Similarly, an ANOVA test was performed to compare the teacher leadership behavior scores among the groups, revealing no significant differences ($F(2, 416) = 1.28$, $p = 28$). These findings indicate successful random assignment among three groups.

Manipulation check

A manipulation check was carried out to ensure the effectiveness of the manipulations. The ANOVA results demonstrated a significant difference in self-esteem scores among the groups ($F(2, 362) = 3.79$, $p = .02$). Specifically, the mean self-esteem score for the high self-esteem group was 4.90 ($SD = 1.57$), for the control group was 4.49 ($SD = 1.59$), and for the low self-esteem group it was 4.38 ($SD = 1.58$). Post-hoc comparisons employing the Tukey HSD test revealed that the high self-esteem group had significantly higher self-esteem scores compared to the low self-esteem group, while the control group also had significantly higher self-esteem scores than the low self-esteem group.

Likewise, an ANOVA test was conducted to examine the positive mood scores across the groups, revealing a significant difference ($F(2, 354) = 5.80$, $p < .01$). The mean positive mood score for the control group was 1.86 (SD =.80), for the high self-esteem group it was 2.21 (SD = 86), and for the low self-esteem group it was 1.94 (SD = .80). Post-hoc comparisons employing the Tukey HSD test indicated that the high self-esteem group had significantly higher positive mood scores compared to the low self-esteem group ($p < .05$). These results confirm the successful manipulation.

The Effect

The study comprised three experimental conditions: a high self-esteem group (HSG), a low self-esteem group (LSG), and a control group. Following the intervention, participants' beliefs in teacher leadership were measured as the dependent variable.

To compare the mean scores of teacher leadership belief (TLB), an analysis of variance (ANOVA) was conducted across the three groups. The results indicated a significant difference in the mean TLB scores ($F(2, 348) = 3.91$, $p < .02$). The mean score for the HSG group was 3.59 ($SD = .88$), for the control group it was 3.34 (SD = .85), and for the LSG group it was 3.31 ($SD = .81$). Post-hoc comparisons using the Tukey HSD test revealed a significant difference between the HSG group ($M = 3.59$, $SD = .88$) and the LSG group ($M = 3.31$, $SD = .81$, $p < .05$). However, no significant differences were found between the control group and the LSG group, or between the control group and the HSG group. These findings support the hypothesis that self-esteem significantly influences teachers' leadership beliefs. Thus, H3 was supported.

General Discussion

The current research aimed to examine the impact of self-esteem on teachers' leadership belief- the extent to which teachers perceive leadership behavior as part of their professional identity. The results have provided experimental evidence that self-esteem significantly enhances teacher leadership beliefs and that teachers who have higher self-esteem are more likely to embrace leadership beliefs. This finding is consistent with the existing literature on the relationship between self-esteem and leadership. Beer et al. (1959), Kipnis and Lane (1962), Hill and Ritchie (1977), Galante and Ward (2017) and Matzler et al. (2015) discovered a positive relationship between self-esteem and leadership. Vašašová et al. (2021) investigated this relationship in the context of teacher leadership and similarly found a positive correlation between self-esteem and teacher leadership. However, none of these studies employed an experimental design and thus could not establish a causal link between self-esteem and leadership. Unlike these studies, the current study provided experimental evidence that self-esteem is a precursor to teacher leadership. In that regard, it made a significant contribution to the extant literature on the antecedents of teacher leadership.

The current study is significant due its experimental design. Podsakoff and Podsakoff (2019) reviewed the methodology of the studies published in general leadership literature during 2015–2018 and found that only 11.5% of studies were experimental. Similarly, Schott et al. (2020) reviewed 90 studies in the educational leadership literature and none of these studies employed experimental designs. Thus, to the best of my knowledge, this is the first study that adopts an experimental investigation into the concept of teacher leadership. By doing so, the study contributes to the field and encourages the integration of experimental designs into future research studies.

The current study tested the effect of self-esteem on teacher leadership "belief" rather than actual "behavior". The future studies could test the impact on actual behaviors. Changing teachers' beliefs could be the first step to encouraging teachers to engage in leadership behavior since teachers' beliefs influence their practices (Fives & Buehl, 2012). Teachers are more likely to embrace new practices which align with their beliefs (Bingham & Hall-Kenyon, 2013). If school leaders can integrate self-esteem workshops into professional development activities and thus enable teachers to adopt leadership beliefs, teachers will be more likely to engage in leadership behaviors.

The study has certain limitations that future studies can address. These areas of improvement could also be valuable suggestions for new studies in the field. First of all, the current study tested the impact of self-esteem on teachers' leadership beliefs. The study results suggested that teachers who have higher self-esteem are more likely to see "leadership" as part of their professional identity, as compared to those who have lower self-esteem. However, there might a discrepancy between beliefs and actual behavior. Therefore, future studies should assess the

impact of self-esteem on actual leadership behaviors rather than just beliefs or perceptions. Secondly, the current study has suggested a causal link between self-esteem and teacher leadership, with self-esteem being a precursor to teacher leadership. However, this relationship could be bi-directional and recursive. In other words, teacher leadership could be leading to the development of self-esteem as well. Future studies need to test the impact of teacher leadership on the improvement of self-esteem. Thirdly, the current study was a brief online experiment. Future studies could involve longitudinal experimental studies where teachers are offered year-long professional development programs and interventions. These studies could assess the impact of self-esteem on both teacher leadership belief and behaviors. Fourthly, the effect of self-esteem was measured immediately after the intervention. Future studies should assess the long-term impact of such interventions. Additionally, an online experiment also could exclude certain demographics. Future studies could be carried out in face-to-face settings and include a more diverse population.

The study has certain practical implications for school leaders. School leaders could include self-esteem workshops into their professional development activities for teachers. These workshops could focus on building or enhancing teachers' self-esteem, which is likely to foster teacher leadership. Secondly, positive relationships enhance self-esteem. Thus, school leaders should build supportive and inclusive school culture. Such a culture develops a sense of belonging and thus increase self-esteem and promotes teacher leadership. Lastly, school leaders need to avoid negative discourse or factors that may affect self-esteem negatively. In sum, by paying attention to self-esteem boosting factors, school leaders can enable every teacher to exercise leadership.

References

Ainscow, M., Barrs, D., and Martin, J. (1998). Taking school improvement into the classroom. *Improving Schools, 1*(3), 43-48.

Angelle, P. S., and DeHart, C. A. (2011). Teacher perceptions of teacher leadership: Examining differences by experience, degree, and position. *NASSP Bulletin, 95*(2), 141-160.

Antonakis, J., Bendahan, S., Jacquart, P., and Lalive, R. (2010). On making causal claims: A review and recommendations. *The Leadership Quarterly, 21*(6), 1086-1120.

Bandura, A. (1997). Self-efficacy: *The exercise of control.* Freeman.

Baumeister, R. F. (1999). Self-concept, self-esteem, and identity.

Baumeister, R. F., Campbell, J. D., Krueger, J. I., and Vohs, K. D. (2003). Does high self-esteem cause better performance, interpersonal success, happiness, or healthier lifestyles? *Psychological Science in The Public Interest, 4*(1), 1-44.

Beer, M., Buckhout, R., Horowitz, M. W., and Levy, S. (1959). Some perceived properties of the differences between leaders and nonleaders. *Journal of Psychology, 47*, 49-56.

Bingham, G. E., & Hall-Kenyon, K. M. (2013). Examining teachers' beliefs about and implementation of a balanced literacy framework. *Journal of Research in Reading, 36*(1), 14-28.

Bolat, O. (2013). *A non-positional teacher leadership approach to school improvement: An action research study in Turkey.* Unpublished doctoral dissertation, The University of Cambridge, Cambridge.

Bolat, O. (2023). Öğretmenlerin liderlik tutumlarına ilişkin bir ölçek geliştirme çalışması. [A study on developing a scale to measure teachers' leadership beliefs] (Under review)

Bolat, Ö., & Antalyalı, Ö. L. (2023). PS Kişisel Eğilimler Envanterinin Psikometrik Özellikleri. [Psychometric Properties of the PS Personal Tendencies Inventory.] *Turkish Studies-Educational Sciences,* 18(2).

Brockner, J. (1988). *Self-esteem at work: Research, theory, and practice.* Heath.

Brown, J. D. (2010). High self-esteem buffers negative feedback: Once more with feeling. *Cognition and Emotion, 24*(8), 1389-1404.

Campbell, J. D. (1990). Self-esteem and clarity of the self-concept. *Journal Of Personality and Social Psychology, 59*(3), 538.

Carrion, R. G., and García-Carrión, R. (2015). Transforming education through teacher leadership. *International Journal of Educational Leadership and Management,* 219-221.

Collinson, D. (2012). Prozac leadership and the limits of positive thinking. *Leadership, 8*(2), 87-107.

Coopersmith, S. (1967). *The antecedents of self-esteem.* H Freeman and Company.

Crocker, J., and Luhtanen, R. (1990). Collective self-esteem and ingroup bias. *Journal Of Personality and Social Psychology, 58*(1), 60.

Diener, E., and Diener, M. (1995). Cross-cultural correlates of life satisfaction and self-esteem. *Journal Of Personality and Social Psychology, 68*(4), 653.

Ding, Z., and Thien, L. M. (2022). Assessing the antecedents and consequences of teacher leadership: A partial least squares analysis. *International Journal of Leadership in Education,* 1-23.

Donnellan, M. B., Trzesniewski, K. H., and Robins, R. W. (2011). *Self-esteem: Enduring issues and controversies.* In T. Chamorro-Premuzic, S. von Stumm, and A. Furnham (Eds.), The Wiley-Blackwell handbook of individual differences (pp. 718–746). Wiley Blackwell.

Durrant, J., and Holden, G. (2006). *Teachers as leaders of learning. Teachers leading change: Doing research for school improvement.* Sage.

Faul, F., Erdfelder, E., Buchner, A., & Lang, A. G. (2009). Statistical power analyses using G* Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods, 41*(4), 1149-1160.

Fives, H., & Buehl, M. M. (2012). Spring cleaning for the "messy" construct of teachers' beliefs: What are they? Which have been examined? What can they tell us?.

Frost, D. (2019). *Teacher leadership and professionality.* In Oxford Research Encyclopedia of Education.

Frost, D., and Durrant, J. (2002). Teachers as leaders: Exploring the impact of teacher-led development work. *School Leadership and Management, 22*(2), 143-161.

Fullan, M. (2007). *Change theory as a force for school improvement. In Intelligent leadership: Constructs for thinking education leaders* (pp. 27-39). Springer Netherlands.

Galante, M., and Ward, R. M. (2017). Female student leaders: An examination of transformational leadership, athletics, and self-esteem. *Personality And Individual Differences, 106*, 157-162.

Gümüş, S., Hallinger, P., Cansoy, R., and Bellibaş, M. Ş. (2021). Instructional leadership in a centralized and competitive educational system: A qualitative meta-synthesis of research from Turkey. *Journal of Educational Administration, 59*(6), 702-720.

Heifetz, R. (2004). *Leadership without easy answers.* Harvard Business School Publishing.

Heifetz, R., and Linsky, M. (2017). *Leadership on the line, with a new preface: Staying alive through the dangers of change.* Harvard Business Press.

Heppner, W. L., and Kernis, M. H. (2011). *High self-esteem: Multiple forms and their outcomes. In Handbook of identity theory and research* (pp. 329-355). Springer New York.

Hill, N. C., and Ritchie, J. B. (1977). The effect of self-esteem on leadership and achievement: A paradigm and a review. *Group and Organization Studies, 2*(4), 491-503.

Holland, R. W., Meertens, R. M., and Van Vugt, M. (2002). Dissonance on the road: Self-esteem as a moderator of internal and external self-justification strategies. *Personality and Social Psychology Bulletin, 28*(12), 1713-1724.

Hu, J., Wang, Z., Liden, R. C., and Sun, J. (2012). The influence of leader core self-evaluation on follower reports of transformational leadership. *The Leadership Quarterly, 23*(5), 860-868.

Huang, T. (2016). Linking the private and public: Teacher leadership and teacher education in reflexive modernity. *European Journal of Teacher Education, 39*(2), 222-237.

Huitt, W. (2004). Self-concept and self-esteem. *Educational Psychology Interactive, 1*, 1-5.

Hunzicker, J. (2017). From teacher to teacher leader: A conceptual model. *International Journal of Teacher Leadership, 8*(2), 1-27.

Ingersoll, R. M., Sirinides, P., and Dougherty, P. (2018). Leadership Matters: Teachers' Roles in School Decision Making and School Performance. *American Educator, 42*(1), 13.

Judge, T. A., and Bono, J. E. (2001). Relationship of core self-evaluations traits—self-esteem, generalized self-efficacy, locus of control, and emotional stability—with job satisfaction and job performance: A meta-analysis. *Journal of Applied Psychology, 86*, 80-92.

Judge, T. A., Erez, A., Bono, J. E., and Thoresen, C. J. (2002). Are measures of self-esteem, neuroticism, locus of control, and generalized self-efficacy indicators of a common core construct? *Journal of Personality and Social Psychology, 83*, 693–710.

Katzenmeyer, M., and Moller, G. (2009). *Awakening the sleeping giant: Helping teachers develop as leaders.* Corwin Press.

Kipnis, D., and Lane, W. P. (1962). Self-confidence and leadership. *Journal of Applied Psychology, 46*(4), 291.

Korman, A. K. (1968). Task success, task popularity, and self-esteem as influences on task liking. *Journal of Applied Psychology, 52*(6p1), 484.

Krieg, S., Smith, K. A., and Davis, K. (2014). Exploring the dance of early childhood educational leadership. *Australasian Journal of Early Childhood, 39*(1), 73-80.

Landis, B., Jachimowicz, J. M., Wang, D. J., and Krause, R. W. (2022). Revisiting extraversion and leadership emergence: A social network churn perspective. *Journal of Personality and Social Psychology.*

Lee-Flynn, S. C., Pomaki, G., Delongis, A., Biesanz, J. C., and Puterman, E. (2011). Daily cognitive appraisals, daily affect, and long-term depressive symptoms: The role of self-esteem and self-concept clarity in the stress process. *Personality and Social Psychology Bulletin, 37*(2), 255–268.

Liljenberg, M. (2016). Teacher leadership modes and practices in a Swedish context–a case study. *School Leadership and Management, 36*(1), 21-40.

Liu, X., Cao, X., and Gao, W. (2022). Does low self-esteem predict anxiety among Chinese college students? *Psychology Research and Behavior Management,* 1481-1487.

Lovett, S. (2018). *Advocacy for teacher leadership: Opportunity, preparation, support, and pathways.* Springer.

Margolis, J., and Huggins, K. S. (2012). Distributed but undefined: New teacher leader roles to change schools. *Journal of School Leadership, 22*(5), 953-981.

Matzler, K., Bauer, F. A., & Mooradian, T. A. (2015). Self-esteem and transformational leadership. *Journal of Managerial Psychology*, *30*(7), 815-831.

Meirink, J., Van Der Want, A., Louws, M., Meijer, P., Oolbekkink-Marchand, H., and Schaap, H. (2020). Beginning teachers' opportunities for enacting informal teacher leadership: Perceptions of teachers and school management staff members. *European Journal of Teacher Education, 43*(2), 243-257.

Mongillo, G., Lawrence, S. A., and Hong, C. E. (2012). Empowering leaders in a master's in literacy program: Teacher candidates' self-efficacy and self-perception as literacy leaders. *Action in Teacher Education, 34*(5-6), 551-565.

Neumerski, C. M. (2012). Rethinking instructional leadership, a review: What do we know about principal, teacher, and coach instructional leadership, and where should we go from here? *Educational Administration Quarterly, 49*(2), 310-347.

Nguyen, D., Harris, A., and Ng, D. (2020). A review of the empirical research on teacher leadership (2003–2017) Evidence, patterns and implications. *Journal Of Educational Administration, 58*(1), 60-80.

OECD (2013), *"The Skills Needed for the 21st Century", in OECD Skills Outlook 2013: First Results from the Survey of Adult Skills,* OECD Publishing

Orth, U., and Robins, R. W. (2014). The development of self-esteem. *Current Directions In Psychological Science, 23*(5), 381-387.

Orth, U., Robins, R. W., and Widaman, K. F. (2012). Life-span development of self-esteem and its effects on important life outcomes. *Journal Of Personality and Social Psychology, 102*(6), 1271.

Pankake, A., and Moller, G. (2007). What the teacher leader needs from the principal. *The Learning Professional, 28*(1), 32.

Pelham, B. W., and Swann, W. B. (1989). From self-conceptions to self-worth: On the sources and structure of global self-esteem. *Journal Of Personality and Social Psychology, 57*(4), 672.

Pierce, J. L., Gardner, D. G., Cummings, L. L., and Dunham, R. B. (1989). Organization-based self-esteem: Construct definition, measurement, and validation. *Academy Of Management Journal, 32*(3), 622-648.

Podsakoff, P. M., & Podsakoff, N. P. (2019). Experimental designs in management and leadership research: Strengths, limitations, and recommendations for improving publishability. The Leadership Quarterly, 30(1), 11-33.

Poekert, P. E. (2012). Learning about teachers learning to lead: Reflections on themes in the teacher leadership literature and their manifestation in a teacher leadership development program in Florida. In *European Conference on Educational Research. Porto, Portugal. Available at https:// www. educ. cam. ac. uk/centres/lfl/about/ PDFs/PoekertECER2014TLPaper. pdf [accessed 20 May 2019]*.

Poekert, P., Alexandrou, A., and Shannon, D. (2016). How teachers become leaders: An internationally validated theoretical model of teacher leadership development. *Research In Post-Compulsory Education, 21*(4), 307-329.

Raghunathan, R., and Trope, Y. (2002). Walking the tightrope between feeling good and being accurate: Mood as a resource in processing persuasive messages. *Journal Of Personality and Social Psychology, 83*(3), 510.

Robins, R. W., Hendin, H. M., and Trzesniewski, K. H. (2001). Measuring global self-esteem: Construct validation of a single-item measure and the Rosenberg Self-Esteem Scale. *Personality And Social Psychology Bulletin, 27*(2), 151-161.

Rogers, J. (2005). Aspiring to leadership—identifying teacher-leaders. *Medical Teacher, 27*(7), 629-633.

Rosenberg, M. (1965). Rosenberg self-esteem scale. *Journal of Religion and Health.*

Schmeichel, B. J., and Vohs, K. (2009). Self-affirmation and self-control: Affirming core values counteracts ego depletion. *Journal Of Personality and Social Psychology, 96*(4), 770.

Schott, C., van Roekel, H., and Tummers, L. G. (2020). Teacher leadership: A systematic review, methodological quality assessment and conceptual framework. *Educational Research Review, 31*, 100352.

Shackelford, T. K. (2001). Self-esteem in marriage. *Personality and Individual Differences, 30*, 371-390.

Shen, J., Wu, H., Reeves, P., Zheng, Y., Ryan, L., & Anderson, D. (2020). The association between teacher leadership and student achievement: A meta-analysis. *Educational Research Review*, *31*, 100357.

Silva, D. Y., Gimbert, B., & Nolan, J. (2000). Sliding the doors: Locking and unlocking possibilities for teacher leadership. *Teachers College Record, 102*(4), 779-804.

Smulyan, L. (2016). Symposium introduction: Stepping into their power: The development of a teacher leadership stance. *Schools, 13*(1), 8-28.

Smylie, M. A., Conley, S., & Marks, H. M. (2002). Section four: Reshaping leadership in action: Exploring new approaches to teacher leadership for school improvement. *Teachers College Record, 104*(9), 162-188.

Sowislo, J. F., & Orth, U. (2013). Does low self-esteem predict depression and anxiety? A meta-analysis of longitudinal studies. *Psychological Bulletin, 139*(1), 213.

Struyve, C., Hannes, K., Meredith, C., Vandecandelaere, M., Gielen, S., & De Fraine, B. (2018). Teacher leadership in practice: Mapping the negotiation of the position of the special educational needs coordinator in schools. *Scandinavian Journal of Educational Research, 62*(5), 701-718.

Sun, J., Chen, X., & Zhang, S. (2017). A review of research evidence on the antecedents of transformational leadership. *Education Sciences, 7*(1), 15.

Supovitz, J., Sirinides, P., & May, H. (2010). How principals and peers influence teaching and learning. *Educational Administration Quarterly, 46*(1), 31-56.

Tafarodi, R. W., & Swann Jr, W. B. (2001). Two-dimensional self-esteem: Theory and measurement. *Personality and Individual Differences, 31*(5), 653-673.

Teacher Leadership Exploratory Consortium. (2011). Teacher leader model standards. Carrboro, NC: Author. Retrieved from http://www. teacherleaderstandards.org/downloads/ TLS_ Brochure.pdf

Trzesniewski, K. H., Donnellan, M. B., and Robins, R. W. (2003). Stability of self-esteem across the life span. *Journal Of Personality and Social Psychology, 84*(1), 205.

Vašašová, Z., Strenáčiková, M., & Žitniaková Gurgová, B. (2021). Teachers' Leadership Behavior in Relation to their Self-Esteem. *The New Educational Review, 66*, 147-156.

Wang, M., & Ho, D. (2020). Making sense of teacher leadership in early childhood education in China. *International Journal of Leadership in Education*, *23*(3), 300-314.

Wenner, J. A., & Campbell, T. (2017). The theoretical and empirical basis of teacher leadership: A review of the literature. *Review of Educational Research, 87*(1), 134-171.

Wojciszke, B., & Struzynska–Kujalowicz, A. (2007). Power influences self–esteem. *Social Cognition, 25*(4), 472-494.

Zeigler-Hill, V., & Showers, C. J. (2007). Self-structure and self-esteem stability: The hidden vulnerability of compartmentalization. *Personality and Social Psychology Bulletin, 33*(2), 143-159.

Zeigler-Hill, V., Besser, A., & King, K. (2011). Contingent self-esteem and anticipated reactions to interpersonal rejection and achievement failure. *Journal of Social and Clinical Psychology, 30*(10), 1069-1096.

This page is intentionally left blank.

www.iejee.com

iejee

International Electronic Journal of Elementary Education
September 2023, Volume 16, Issue 1, 111-122

The Impact of Brain-Based Learning on Students' Intrinsic Motivation to Learn and Perform in Mathematics: A Neuroscientific Study in School Psychology

Amjad Islam Amjad[a,*], Musarrat Habib[b], Umaira Tabassum[c], Gulshan Fatima Alvi[d], Naveed Ahmad Taseer[e], Iqra Noreen[f]

Received : 25 May 2023
Revised : 31 July 2023
Accepted : 29 September 2023
DOI : 10.26822/iejee.2023.318

[a,*] **Corresponding Author:** Amjad Islam Amjad, School Education Department, Punjab, Pakistan.
E-mail: amjad_14@yahoo.com
ORCID: https://orcid.org/0000-0002-4250-7526

[b] Musarrat Habib, Department of Education, The University of Lahore, Lahore, Pakistan.
E-mail: musarrat.javaid@ed.uol.edu.pk

[c] Umaira Tabassum, Department of Education, Guangzhou University, Guangzhou, China.
E-mail: umairatabassumsattar@gmail.com

[d] Gulshan Fatima Alvi, Department of Education, Lahore Leads University, Lahore, Pakistan.
E-mail: drgulshan.edu@leads.edu.pk

[e] Naveed Ahmad Taseer, Institute of Education and Research, University of the Punjab, Lahore, Pakistan.
E-mail: taseer642nat1@gmail.com

[e] Iqra Noreen, Institute of Education and Research, University of the Punjab, Lahore, Pakistan.
E-mail: Iqrapu1@gmail.com

Abstract

The current study aimed to explore the effect of Brain-Based Learning on students' intrinsic motivation (IM) to learn and perform in mathematics. Owing to the educational implications of Neuroscience, the researchers planned the mixed-methods experimental study with a convergent parallel research design. The participants were eighth-graders enrolled in a boy's public secondary school in the district Kasur of Pakistan in the academic year 2021-2022. Students were taught mathematics using the traditional lecture method in the baseline phase (A) and withdrawal phase (A). In contrast, in the treatment phase (B), they were taught mathematics with activities based on the BBL approaches and principles. The mathematics motivation scale and observation (field notes) were used to collect quantitative and qualitative data simultaneously. Data were analysed using one-way repeated measure ANOVA and thematic analysis. The researchers revealed that BBL significantly affects students' IM to learn and perform in mathematics. After qualitative analysis, it was found that visual story-telling, role-playing, i-Think maps, back-to-board, and Kick Me Poison Box are the most effective BBL-based activities which keep them engaged, provide emotional support and contribute to students' IM to learn and perform in mathematics. The teachers were suggested to develop BBL-based activities to teach mathematics and to influence students' psychology in the schools.

Keywords:

Neuroscience, Brain-Based Learning, School Psychology, Intrinsic Motivation, Visual Story-Telling

Introduction

Neuroscience with its relevance and functions has emerged as a powerful field of research with immense potential to revolutionise various disciplines (Amjad et al., 2022a; Palser et al., 2022). School psychology is one such field that benefits significantly from integrating neuroscience (Baker et al., 2021; Benson et al., 2019). As our understanding of the brain and its cognitive processes expands, educators and school psychologists have begun

to recognise neuroscience's profound implications on enhancing student learning, well-being, and overall educational outcomes (Tortella et al., 2021). This research paper aimed to explore the impact of neuroscience (Brain-Based Learning) in the field of school psychology, examining how the application of BBL principles can inform and improve educational practices. By integrating neuroscience findings into instruction, intervention, and assessment, school psychologists can enhance their ability to address the diverse needs of students and create more effective learning environments.

The objective of maximising students' cognitive, social, and emotional growth is at the heart of educational psychology (Immordino-Yang et al., 2019). Brain-Based Learning (BBL) provides valuable insights into the underlying neural mechanisms that govern these processes (Jailani, 2021). By understanding the brain's structure, functioning, and flexibility, school psychologists can introduce interventions and educational strategies that align with the biological foundations of learning and development (Wilcox et al., 2021). The BBL approach can greatly benefit intervention strategies to support students' learning and well-being (Baratali & Zardeini, 2023). By identifying the neural processes associated with specific learning difficulties or behavioural challenges, school psychologists can provide interventions to target these underlying neural mechanisms (Sanetti & Collier-Meek, 2019). The BBL offers valuable insights into how children's brains acquire, process, and retain information (Li & Lan, 2022). Understanding brain functions associated with attention, memory, motivation, and emotional regulation can guide educators in designing instructional strategies that capitalise on these processes (Drigas & Karyotaki, 2019).

From elementary school through university, mathematics is now an integral course for students (Oljayevna & Shavkatovna, 2020). In the Pakistani school system, it is considered a compulsory subject (Amjad et al., 2022b). Most students think that mathematics is complicated for students since it requires them to learn formulas (Das, 2019), carefully follow examples (Udjaja et al., 2018), and work with sometimes less interesting subject matter (Appelgate & Jurgenson, 2022). Although memorising extensive mathematical notations is a common perception of mathematical ability (Pascual, 2022), Nematillayevna (2021) argues that what matters is learning how to apply those notations effectively when faced with a challenge.

Mathematics significantly impacts our personal and professional lives, and its essential for individual and national economic development (Maass et al., 2019). Its multidisciplinary nature aids a country's scientific and technological progress (Bano et al., 2018). Mathematics supports engineering, physics, sociology, chemistry, and arts (Weeden et al., 2020). It profoundly impacts scientific and technological advancements, from research to products and business methods (Engelbrecht et al., 2020). Given its broad influence, mathematics is vital in modern curricula (Alayont, 2022). Tokac et al. (2019) recommend providing students with resources through mathematics education to succeed in today's advanced nations.

According to Khan et al. (2020), Pakistani students show less performance in mathematics as compared to other school-level subjects. Overall, it was also found that students' arithmetic ability to perform in mathematics was not good, which led to decreasing motivation to learn and perform in mathematics. It was also found that students between the ages of four and eight perform poorly globally in mathematics. According to Stojanović et al. (2021), teachers' ability to teach, the calibre of the pupils, and their motivation all play a significant role in how well students achieve in mathematics. Additionally, it is clear from the literature that students struggle academically in mathematics classes because they lack motivation.

Influence of BBL on Learning and Motivation

The BBL is based on the principles of Neuroscience, which was first used in American schools in the 1990s (Ferreira & Rodríguez, 2022). From the tiniest cellular level to the greatest cerebral circuits, neuroscience explores how the brain works to learn and remember things (Glaser et al., 2019). Teachers must understand how to teach using what scientists have discovered about the brain to learn in a way compatible with their brain. Applying the ideas and methods found through research on the brain is necessary for teaching in a way that is in harmony with the brain (Ferreira & Rodríguez, 2022). Understanding how the brain functions, what influences the brain, and the educational consequences of these insights have changed over time (Tan & Amiel, 2022). One of the BBL's founders, Caine and Caine (1991), stated the following approaches based on 12 BBL principles.

Relaxed Alertness: Triana and Zubainur (2019) suggest that a peaceful and attentive atmosphere should be provided when students are given a difficult task. For successful problem-solving, students must be relaxed and engaged. Saleh and Mazlan (2019) argue that a challenging learning atmosphere with minimal physical risk is essential to ensure students learning.

Orchestrated immersion: It is related to immerse students physically, mentally, and emotionally in a topic to support learning (Saleh & Mazlan, 2019). Triana and Zubainur (2019) argue that students can use their methods to solve problems, helping them remember the principles. Critical thinking and acquiring new information are essential for learning.

Active processing of experience: To foster learning, Triana and Zubainur (2019) suggest creating environments like small discussion groups that motivate students to engage with the material. Saleh and Mazlan (2019) added that it enables students to internalise information, value it, make connections, and form relevant opinions or decisions.

Motivation for Mathematics Learning

Several factors and types of motivation affect mathematics learning motivation (Hung et al., 2019). Wilkie and Sullivan (2018) suggest techniques to influence internal and external motivation. IM happens when a person is motivated to succeed by themselves. Extrinsic motivation (EM) is external and easier to cultivate for mathematics study than IM. External sources like instructors' motivation and peers' approval of a positive learning environment can strengthen EM. IM is more difficult to develop. El-Adl and Alkharusi (2020) argue that IM leads to self-satisfaction as students strive for learning success. IM happens when students find personal value in something and recognise its potential benefits. Weidinger et al. (2017) found that IM rises at the start of school but decreases in the early years. Some students experience a more significant drop than others, while some show no decrease. Learners' IM is based on meeting their needs for competence, instruction, and teaching activities (Ryan & Deci, 2016). Good grades should motivate students to learn, while poor grades can reduce their sense of intellectual ability and IM. Corpus and Wormington (2014) found that lower scores in elementary school correlated with lower IM.

BBL and Students' Intrinsic Motivation

Neuroscience-based teaching approach (BBL) have educational applications and have been used to study student motivation (Al-Balushi & Al-Balushi, 2018), academic performance (Amjad et al., 2022b), and problem-solving skills (Pohan et al., 2020). Mekarina and Ningsih (2017) found that BBL is an effective teaching approach for mathematics and can enhance students' motivation and performance. BBL helps students use the right part of the brain to learn and perform in mathematics. This enables them to make the most of their brains to understand mathematics concepts. Yu and Singh (2018) studied the relationship between teachers' teaching strategies and students' mathematics learning motivation. SEM analysis revealed a positive relationship between teachers' teaching strategies, self-efficacy, and motivation.

BBL and students' IM to learn and perform in mathematics are interconnected concepts with great promise for transforming the educational landscape. BBL emphasises aligning teaching strategies with the brain's natural learning processes (AlAkayleh & Al-Zoubi, 2023), while IM refers to students' internal drive and interest in learning (Bai et al., 2021). When BBL principles are applied effectively in mathematics instruction, they can intensely impact students' IM (Solihatin & Syahrial, 2019). Here are some key points to explore the relationship between BBL and IM in mathematics:

Engaging Learning Experiences: BBL promotes active and experiential learning, meaningfully allowing students to explore and engage with mathematical concepts. Students' curiosity is piqued by incorporating hands-on activities, real-world problem-solving, and interactive experiences, fostering a more profound interest and IM to delve into the subject (El-Adl & Saad, 2019).

Personalised Learning: BBL recognises the uniqueness of each student's brain and learning style. By tailoring instruction to individual needs and preferences, students feel more autonomy and ownership over their learning journey. This sense of control and personalisation can ignite IM as students' experience and as a sense of accomplishment in mastering mathematical challenges (Baratali & Zardeini, 2023).

Positive Emotional Climate: BBL has shown that emotions play a crucial role in learning and memory. A brain-based approach to mathematics instruction creates a positive and supportive learning environment where students feel safe to take risks, make mistakes, and embrace challenges. When students experience positive emotions in the classroom, it enhances their IM to learn and perform in mathematics (Baratali & Zardeini, 2023).

Meaningful Relevance: BBL advocates connecting new knowledge and existing experiences. By demonstrating the practical relevance of mathematics in everyday life, careers, and various fields of study, students can see the significance of learning mathematics, increasing their IM to learn and perform in the subject (Apeh & Iyiegbuniwe, 2021).

Mastery and Growth Mindset: BBL supports the idea of a mastery and growth mindset. Students are encouraged to see effort and perseverance as the path to improvement rather than fixating on grades or performance. This mindset fosters IM as students become more focused on the joy of learning and the satisfaction of personal development rather than external rewards (Hofer, 2022).

Challenging but Manageable Tasks: BBL advocates for presenting challenges that are neither easy nor difficult. By providing students with appropriately challenging mathematical tasks, they experience a state of flow where they are fully engaged and immersed in the learning process. This flow state enhances students' IM and fosters a love for mathematics (Apeh & Iyiegbuniwe, 2021).

Research Objective

Many researchers in Pakistan have employed different teaching strategies to increase students' motivation to learn mathematics (Farooq et al., 2020; Talpur et al., 2021). Why pupils fail to acquire higher grades at all levels, particularly at the elementary level, is still a concern (Mushtaq, 2021). The authors, in the present mixed-methods' study investigated how the BBL affects eighth graders' IM to learn and perform in mathematics. The researchers developed a null hypothesis based on the study aim that BBL had no significant effect on elementary-level pupils' IM to learn and perform in mathematics. The researchers also developed a research question to explore which teaching activities could enhance students' IM to learn and perform in mathematics at the elementary level.

Methods

The authors intended to study the influence of BBL-based intervention on students' IM to learn and perform in mathematics on a deeper level. Therefore, an in-depth study of the phenomenon was designed using mixed-methods research. The pragmatic research philosophy guided the present study. The problem under investigation was explored using quantitative and qualitative approaches.

Research Design

The research design of the study was convergent parallel. In this design, researchers collected quantitative and qualitative data simultaneously and independently. The results from both data sets were then compared and contrasted to identify any convergence or differences. The researchers used an experimental-based quantitative investigation during the different phases of the study. It was an 18-week experimental study that had three different stages. The experimental phase was carried out to investigate the change in applied behaviour analysis (A-B-A) in students' IM to learn and perform in mathematics. It was carried out in three stages of equal length (6 weeks). In the first stage A (baseline phase), students were taught eighth-grade mathematics using the traditional lecture method. In the second stage B (treatment phase), the students were taught mathematics using activities based on BBL principles. In the third stage of study A (withdrawal phase), the researchers withdrew the intervention, and students were again taught mathematics using the traditional lecture method.

Study Content

In the present mixed-methods investigation, elementary-level students were taught nine units overall. In the baseline phase, they were taught units (1-3) named "Operation on Sets," "Real Numbers," and "Number System." During the treatment phase, they were taught units (4-6) named "Financial Arithmetic," "Polynomials," and "Factorization, Simultaneous Equations." In the withdrawal phase, they were taught the three units (7-9) named "Fundamentals of Geometry," "Practical Geometry," and "Areas of Volumes.". It is evident that public school teachers typically teach mathematics with traditional lecture method using whiteboard practice only. Based on the nature of lower IM among Pakistani students to learn and perform in mathematics, the researchers in the current study deployed BBL-based teaching in the treatment phase.

Participants

The participants of the current mixed-methods study were elementary-level students. These students were considered based on the argument that elementary-level students show less motivation to learn and perform in mathematics. The researchers selected a school using convenience sampling and selected all the eighth-grade students enrolled in a public secondary school in the Kasur district. The total number of students in eighth grade was 39, who were taught in a single class section as per the standardised student-teacher ratio (40/1). Single class section was selected as study sample and single-subject research was designed to measure the influence of BBL-based teaching on students' IM to learn and perform in mathematics.

Data Collection Tools

The researchers used two data collection tools simultaneously to collect data in the present study. The authors developed an observation guide to take field notes on students' IM in the class. The observation guide was designed using the different aspects related to the students' IM. To collect quantitative data, Mathematics Motivation Scale (MMS) developed by Zakariya and Massimiliano (2021) was used during all study phases. It was developed on a five-point Likert scale with options ranging from strongly disagree (1) to strongly agree (5). It had different items to measure students' IM to learn and perform in mathematics. The IM of the study respondents was measured using quantitative items like "If I can, I want to get better grades in this class than most of the other students." The reliability of the MMS was ensured using Cronbach's Alpha, and it ranged from .73 to .86, which was satisfactory to use, according to Hair et al. (2021).

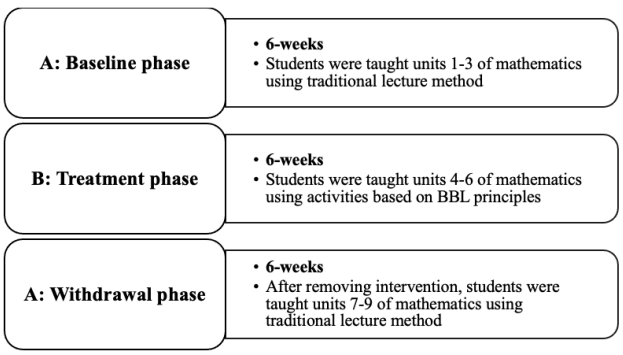
Procedure

The researchers deployed mixed-methods approaches in the current single-subject A-B-A experimental study. It was carried out in three phases. In the first stage A (baseline), the researchers taught the first three units of eighth-grade mathematics with

the traditional lecture method and measured their IM to learn and perform in mathematics three times using the MMS at the regular interval of two weeks. Meanwhile, the researchers also took field notes according to the observation guide. In the treatment phase (B), the researchers taught units 4-6 with the help of activities based on the BBL principles. To get the BBL intervention's effect, all the activities were designed by following the BBL approaches; a) relaxed alertness, b) orchestrated immersion, and c) active processing and 12 principles of BBL. The researchers devised brainstorming activities to improve class engagement and students' social and emotional involvement. They were given various mathematical problems to solve and discuss with their classmates. Students were assigned different roles to play in role-playing activities to improve their understanding of various concepts related to financial mathematics, polynomials, and factorisation, simultaneous equations. In one of the exercises, students were given the roles of judge, lawyer, and petitioner for the division of inherited property according to Islamic laws after the death of a father who left behind three sons, a daughter, and a widow. In the following units, students were given different maps, sketches, tree maps, and flowcharts to think, relate, and identify the correct solutions for teaching ideas linked to factorisation and other functions of algebraic expressions. The researchers devised activities based on the students' visual imagery to teach distinct views from the unit covering topics concerning simultaneous linear equations. Short videos, simulations, and images about simultaneous linear equations were shown to them. During the intervention phase (B), the researchers measured students' IM to learn and perform in mathematics at the regular interval of two weeks. The authors also took field notes after observing their engagement and participation in the activities based on BBL principles. In the withdrawal phase A, the researchers withdrew the BBL-based intervention. Students were again taught with the lecture method. The authors took field notes for qualitative data, while quantitative data was measured using MMS at regular intervals of two weeks.

Figure 1

Phase-Wise Study Progression



In Figure 1, the researchers presented the phase-wise study progression. During each stage, three measurements for IM to learn and perform in mathematics were taken using MMS and filed notes.

BBL's founders, Caine and Caine (1991), presented 12 BBL principles. The authors developed teaching activities for treatment phase (B) in this investigation based on the 12 principles. The detail of the activities is presented in Table 1.

Ethical Considerations

In this experimental investigation, the researchers meticulously adhered to established research ethics for the adolescent respondents. Informed consent was obtained from all participating respondents and their parents, signifying their voluntary agreement to participate in the study. After recruitment into the investigation, the participants underwent a comprehensive briefing, acquainting them with the research procedures, study duration, and the prescribed code of ethics. All participants were assured of anonymity to protect their personal information, as neither their names nor locations were disclosed, mitigating any risk of confidentiality breaches. The experiment took place in natural settings, devoid of any interventions or activities that could potentially cause harm to the participants. The activities conducted during the study were thoughtfully aligned with the prevailing school culture and education calendar, ensuring seamless integration into the participants' daily routines. Each activity was precisely organised to fit within a time frame of 45 minutes. Considering linguistic diversity, a bilingual version of the research instrument, comprising both English and Urdu, was employed for quantitative data collection. This thorough modification made it easier for the participants to understand each item and provide meaningful and precise responses.

Data Analysis

In this paper, the authors deployed MMS to collect quantitative data and field notes to collect students' qualitative responses during an 18-week study. The quantitative data was analysed using the SPSS (Version 26) software, and qualitative data were analysed using the NVivo (Version 12 pro) software. The researchers deployed one-way repeated measure ANOVA to find the influence of BBL-based instruction on students' IM to learn and perform in mathematics. The qualitative data was analysed using the thematic data analysis technique by following the six-stage guidelines of Braun and Clarke (2006). The interpretation of the analysis is presented below.

Results

After the data analysis, quantitative and qualitative

Table 1

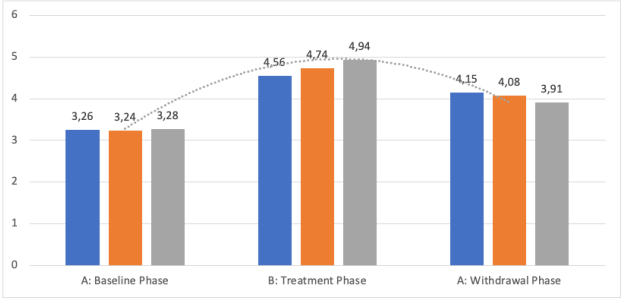
List of Activities and Features Based on BBL Principles

Treatment/Activity	Features	BBL Principles
Visual Story-telling	• Visual imagery • Picture metaphor	• The brain understands and remembers best when facts and skills are embedded in natural spatial memory. • Learning always involves conscious and unconscious processes. • Appropriate environment, music, and aroma excite rain activity. • Each brain is unique.
Role Play	• Emotion in learning • Kinesthetic • Auditory	• A positive climate stimulates brain function. • Learning is enhanced by challenges and inhibited by threats. • Learning engages whole physiology. • Each brain is unique.
Back to Board	• Kinesthetic • Brainstorming • Verbal	• The search for meaning is innate. • The search for meaning comes through the brain patterning process. • Learning always involves conscious and unconscious processes. • Learning engages whole physiology. • Complex and active experiences involving movement stimulate brain development. • Each brain is unique.
Kick Me Poison Box	• Making connection • Develop meaning • Thinking through analogies	• The search for meaning is innate. • Learning always involves conscious and unconscious processes. • Learning always takes place in two memory approaches, retaining facts, skills, and procedures or making sense of experience. • The brain can quickly grasp and remember facts and skills embedded in its memory space. • Each brain is unique.
i-THINK Map	• Brainstorming • Visual imagery • Demonstration of students' understanding	• The brain is unique and a parallel processor. • Learning always takes place in two memory approaches, retaining facts, skills, and procedures or making sense of experience.

data results are presented and interpreted below. The researchers collected data nine times during the current mixed-methods longitudinal experimental study. The mean scores for students' IM to learn and perform in mathematics are presented below.

Figure 2

Effect of BBL on Students' IM to Learn and Perform in Mathematics



In Figure 2, the researchers presented the results of quantitative data to explore the effect of BBL on students' IM to learn and perform mathematics during the study phases. In the baseline phase (*A*), when students were taught with the traditional lecture method, their scores (M_1 = 3.26, M_2 = 3.24, M_3 = 3.29) were lesser than the scores during the treatment phase (*B*) (M_4 = 4.56, M_5 = 4.74, M_6 = 4.94) when they were taught mathematics using the activities bases on BBL principles. In the withdrawal phase (*A*), when the researchers withdrew from the BBL-based intervention, the scores (M_7 = 4.15, M_8 = 4.08, M_9 = 3.91) showed a decline from the treatment phase (*B*). It was found that the increase in mean scores in the treatment phase (*B*) and the decrease in the mean score during the withdrawal phase (*A*) were owed to the BBL-based intervention.

A one-way repeated measure ANOVA was utilised to test the null hypothesis. The mean scores of distinct phases within the study were examined at a significance level of .001. This analysis assumes homogeneity of variance among the differences observed among study phases. The results of the assumption of sphericity are displayed in Table 2.

Table 2 displays the results of Mauchly's test, indicating a significant violation of the assumption of sphericity (Mauchly's W = .000 < .001). In such cases, the adjusted sphericity procedure is warranted. Given that the value of adjusted sphericity exceeds .75, Huynh-Feldt's correction was employed to interpret the outcomes of the BBL on students' IM to learn and perform in mathematics below.

The assumption of sphericity was assessed and found to be 132.71, higher than the value of .75. Consequently, Huynh-Feldt's correction was applied to analyse the differences in mean values for students' IM across nine measurements within three phases. The impact of BBL on eighth-graders' IM during the intervention phase was examined, and post-Huynh-Feldt's correction, the sphericity-adjusted F-statistic was found to be $F(2.95) = 111$, with a p-value of 0.000 ($p < .001$). The effect size, represented by partial eta squared (ηp2), was calculated to be .75, signifying a statistically significant impact of the BBL activities on students' IM to learn mathematics, meeting the criteria set by Latoszek (2020) for statistical significance. Consequently, it can be inferred that the BBL-based activities, including visual imagery, role play, i-Think map, and brainstorming, had a significant effect on students' IM to learn and perform in mathematics throughout the treatment phase (B), specifically in learning financial arithmetic, polynomials, and factorisation related topics.

In this part of the results, the researchers presented the analysis for the qualitative data. The researchers took field notes daily during all the study phases by following the observation guidelines. After applying thematic data analysis, the authors found the following themes.

Class Engagement and Interaction

The researchers found that when the students were taught using the traditional lecture method, students showed poor class management in the baseline phase. They were less interactive and most of the time passive in class. It was also found that few of the students never asked a question to the teacher or peers. While in the treatment phase, they were more engaged. Perhaps, it was the effect of activities designed based on BBL principles. In this phase, students showed enthusiasm and interest, and quantitative results showed that they got the highest IM to learn and perform in mathematics during this phase. In the withdrawal phase, the researchers observed that students' class engagement and interaction decreased from the baseline phase. Although it remained better than the baseline phase, it significantly reduced from the treatment phase. When the authors matched the data triangulation results, it was concluded that both types of data support each other and validate the influence of BBL-based teaching on students' IM to learn and perform

Table 2

Mauchly's Test of Sphericity

Within Subjects Effect	Mauchly's W	Approx. Chi-Square	df	Sig.	Epsilon		
					Greenhouse-Geisser	Huynh-Feldt	Lower-bound
Intrinsic motivation	.000	288.775	35	.000	.340	.369	.125

Table 3

Mean and Standard Deviation Values and Tests of Within-Subjects Effects

Variable	M	SD	N	df	Sphericity assumed	*Effect*	*F ratio*	Sig.	Partial Eta Squared
Week 2	3.26	.607	39	2.95	132.71	IM	111	.000	.75
Week 4	3.24	.669	39						
Week 6	3.29	.648	39						
Week 8	4.56	.348	39						
Week 10	4.74	.275	39						
Week 12	4.94	.125	39						
Week 14	4.15	.420	39						
Week 16	4.08	.427	39						
Week 18	3.91	.624	39						

in mathematics. It showed that activities such as visual story-telling, role-playing, i-Think maps, back-to-board, and Kick Me Poison Box effectively increased students' IM during the treatment phase.

Emotional Support

The authors found that the emotional support of students was lowest during the baseline phase, and it enhanced dramatically during the treatment phase. In the withdrawal phase, it was although better than the baseline phase but significantly decreased from the treatment phase. This fluctuation of emotional support in class was owed to implementing BBL-based teaching. The teacher taught them such activities, which were designed for the BBL principles, and these principles focus on emotional attachment. It was found that when they were part of the groups' activities, they showed an empathetic attitude toward each other. After triangulating the data for validation, it was investigated that activities such as role-playing, visual story-telling, i-Think maps, and back-to-board are effective activities for students' emotional support and contributed to increasing students' IM during the treatment phase.

Discussion

The present mixed-methods study aimed to examine the impact of BBL on students' IM to learn and perform in mathematics. During the research, participants were exposed to traditional lecture-based instruction in the baseline and withdrawal phases, while the treatment phase entailed BBL-based activities and principles. The study's results revealed a significant enhancement in students' IM when exposed to the BBL approach, attributing this improvement to the intervention during the treatment phase. Concerns over enhancing students' motivation in mathematics have long occupied educators and stakeholders (Farooq et al., 2020; Mushtaq, 2021; Talpur et al., 2021), revealing the positive impact of the BBL teaching approach on students' IM in mathematics, particularly significant. Wilkie and Sullivan (2018) posited that IM plays a pivotal role in overall motivation, strengthening efforts toward achieving academic goals, especially in mathematics.

The literature strongly indicates that students' inadequate motivation contributes to a high rate of mathematics failure (Heyder et al., 2020). In this context, the current study offers a promising explanation by proposing BBL-based activities to enhance students' IM and, consequently, their likelihood of academic success. Froiland and Worrell (2016) support this claim, highlighting that intrinsically motivated students outclass in various educational events, displaying sharp passion, challenge-seeking tendencies, greater participation, and enhanced performance. The thematic analysis also supports the argument that when students were taught mathematics with BBL-based activities, they felt more passion and energy to solve mathematical problems.

An essential consideration is the students' classroom engagement, with motivation being critical. Motivated students actively contribute to class work, enthusiastically engaging in tasks and fulfilling deadlines (Musu-Gillette et al., 2015). Embracing BBL-based activities in mathematics instruction can encourage students' motivation by addressing their needs in terms of competency, teaching, and pedagogical experiences, thereby fostering academic achievement and the attainment of favourable grades. Qualitative analysis showed that when BBL-based interventions were applied in the class, it enhanced their class engagement and participation. It further supported the quantitative findings, significantly increasing students' IM during treatment.

Ryan and Deci (2016) reinforce the relationship between motivation and academic achievement, asserting that successful educational outcomes are a powerful motivator for students. Conversely, lower scores may undermine students' confidence in their intellectual abilities and lead to demotivation. In this context, BBL emerges as a valuable approach to nurturing motivation, enabling students to achieve higher scores and find fulfilment in their academic pursuits. Thus, it can be argued that BBL is an effective approach with the potential to enhance students' IM and engage them in class participation. Their passion, energy, and interest in solving mathematical problems increased when they were taught with the help of BBL-based activities.

The research findings, determined through one-way repeated measure ANOVA, conclusively demonstrate that BBL-based activities significantly influence students' IM to learn and perform in mathematics. These results validate the findings of Mekarina and Ningsih (2017), who support the efficacy of BBL as a teaching strategy, enhancing both motivation and mathematical proficiency among students. Furthermore, the estimated marginal mean score highlights a positive relationship between the BBL-based teaching approach and students' IM to learn and perform in mathematics, supporting the study findings of Yu and Singh (2018) in their investigation of the interplay between teachers' strategies and students' motivation for mathematics learning. Additionally, Effendi and Marlina (2021) emphasise the potential of the BBL model to elevate students' motivation in mathematical communication, aligning with the present study's outcomes.

Thus, this research digs into the field of educational exploration, revealing the remarkable potential of BBL to ignite students' IM in the pursuit of mathematical

learning and performance. The findings emphasise the significance of incorporating the BBL approach in mathematics education to unlock students' full potential and pave the way for academic excellence. As educators and stakeholders embark on the path ahead, may these insights serve as a guiding light to invigorate the hearts and minds of future learners, fostering a generation of motivated individuals poised for success in their mathematical endeavours. Hence, it can be argued that the implications of neuroscience in school can effectively affect students' psychology.

Conclusion

The authors performed the convergent parallel study using experimental conditions. Students overcame the fear and concerns about enhancing students' IM to learn and perform in mathematics which has challenged teachers over the decades. Academia also believes that students' IM is one of the crucial aspects for students' success in school psychology. The current study was designed to investigate the effect of BBL on students' IM to learn and perform mathematics at the elementary level. During the intervention phase of the study, the researchers developed mathematics learning activities based on the approaches and principles of BBL. The visual analysis presented a comparison in all stages of the study, which revealed that the increase in IM during the treatment phase (B) was due to the intervention provided to students for mathematics learning. The decline in IM during the withdrawal phase (A) further supported the study's findings that the results started decreasing when the intervention was withdrawn. The qualitative data also supported the quantitative data findings when data triangulation was performed. Thus, the current study concluded that BBL significantly affects students' IM to learn and perform in mathematics.

Limitations, Study Implication, and Future Research

The present investigation had certain limitations. Being a mixed-methods study, the researchers employed an experimental setting on a single subject (A-B-A) with a relatively smaller sample size and confined to a specific region encompassing one school and class. While the A-B-A design effectively addresses potential threats like history and maturation by equalising the duration of all study phases, concerns arise regarding the generalizability of the results to a broader context. Moreover, the repetitive use of the same measurement tool (MMS) for data collection at two-week intervals might introduce potential biases in communicating results. Additionally, the study's budget constraints necessitated the creation of BBL-based activities using cost-effective, non-elaborate materials to foster a conducive environment for students' engagement with mathematical concepts. These limitations could potentially restrict the scope and applicability of the current study.

Despite these limitations, the study's findings have practical implications for students and teachers at the elementary school level. Adopting a classroom environment that emphasises BBL may foster a sense of comfort and ease for teachers and students, potentially alleviating the stress associated with achieving higher academic scores. The investigation demonstrated a significant increase in students' IM to comprehend and apply mathematical concepts, highlighting the potential of BBL as a valuable approach for enhancing students' class participation, engagement, and retention.

While the current study provides valuable insights, it also suggests further exploration. Future researchers may consider employing only a qualitative approach to thoroughly assess the impact of BBL-based activities on students' IM in learning and performing mathematics. They can deploy semi-structured interviews or focused group discussions to explore teachers' and students' dispositions and applications of BBL-based interventions. The current findings could be replicated by designing experimental studies utilising alternative designs beyond the A-B-A. Such endeavours would contribute to a more comprehensive understanding of BBL strategies' potential benefits and effectiveness in mathematics education.

References

AlAkayleh, S. S., & Al-Zoubi, A. M. (2023). The Effectiveness of a Program Based on Brain-Based Learning in Acquiring Mathematical Skills among Students with Learning Difficulties in Math. *Resmilitaris*, *13*(3), 3076-3082.

Alayont, F. (2022). A Case for Ethics in the Mathematics Major Curriculum. *Journal of Humanistic Mathematics, 12*(2), 160-177. https://doi.org/10.5642/jhummath.CXSI3022

Al-Balushi, K. A., & Al-Balushi, S. M. (2018). Effectiveness of Brain-Based Learning for Grade Eight Students' Direct and Postponed Retention in Science. *International Journal of Instruction, 11*(3), 525-538. https://doi.org/10.12973/iji.2018.11336a

Amjad, A. I., Habib, M., & Saeed, M. (2022b). Effect of brain-based learning on students' mathematics performance at the elementary level. *Pakistan Journal of Social Research, 4*(03), 38-51. https://doi.org/10.52567/pjsr.v4i03.684

Amjad, A. I., Tabbasam, U., & Abbas, N. (2022a). The Effect of Brain-Based Learning on Students' Self-Efficacy to Learn and Perform Mathematics: Implication of Neuroscience into School Psychology. *Pakistan Languages and Humanities Review*, *6*(3), 683-695.

Apeh, H. A., & Iyiegbuniwe, O. A. (2021). Effects of brain-based learning strategies on secondary school students' motivation to learn in Federal Capital Territory, Abuja, Nigeria. *International Journal of Research*, *10*(5), 19-32. https://doi.org/10.5861/ijrse.2021.5057

Appelgate, M. H., & Jurgenson, K. (2022). How engagement with mathematics in an integrated STEM lesson evolved over four years. *Investigations in Mathematics Learning, 14*(1), 63-86. https://doi.org/10.1080/19477503.2021.2023965

Bai, S., Hew, K. F., Sailer, M., & Jia, C. (2021). From top to bottom: How positions on different types of leaderboards may affect fully online student learning performance, intrinsic motivation, and course engagement. *Computers & Education*, *173*, 104297. https://doi.org/10.1016/j.compedu.2021.104297

Baker, D. H., Vilidaite, G., Lygo, F. A., Smith, A. K., Flack, T. R., Gouws, A. D., & Andrews, T. J. (2021). Power contours: Optimising sample size and precision in experimental psychology and human neuroscience. *Psychological Methods*, *26*(3), 295. https://doi.org/10.1037/met0000337

Bano, M., Zowghi, D., Kearney, M., Schuck, S., & Aubusson, P. (2018). Mobile learning for science and mathematics school education: A systematic review of empirical evidence. *Computers & Education, 121*, 30-58. https://doi.org/10.1016/j.compedu.2018.02.006

Baratali, M., & Zardeini, A. Z. (2023). Identification of curriculum objectives of brain-based education and positive education. *Journal of Fundamentals of Mental Health*, *25*(2), 69-79.

Benson, N. F., Floyd, R. G., Kranzler, J. H., Eckert, T. L., Fefer, S. A., & Morgan, G. B. (2019). Test use and assessment practices of school psychologists in the United States: Findings from the 2017 National Survey. *Journal of School Psychology*, *72*, 29-48. https://doi.org/10.1016/j.jsp.2018.12.004

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, *3*(2), 77-101. https://doi.org/10.1191/1478088706qp063oa

Caine, R. N., & Caine, G. (1991). Making connections: Teaching and the human brain. ERIC.

Corpus, J. H., & Wormington, S. V. (2014). Profiles of intrinsic and extrinsic motivations in elementary school: A longitudinal analysis. *The Journal of Experimental Education, 82*(4), 480-501. https://doi.org/10.1080/00220973.2013.876225

Das, K. (2019). Role of ICT for Better Mathematics Teaching. *Shanlax International Journal of Education*, *7*(4), 19-28. https://doi.org/10.34293/education.v7i4.641

Drigas, A., & Karyotaki, M. (2019). Executive Functioning and Problem Solving: A Bidirectional Relation. *Int. J. Eng. Pedagog.*, *9*(3), 76-98. https://doi.org/10.3991/ijep.v9i3.10186

Effendi, K. N. S., & Marlina, R. (2021). The effect of motivation towards mathematical communication in mathematics learning with brain-based learning model. *Jurnal Program Studi Pendidikan Matematika, 10*(2), 808-819. https://doi.org/10.24127/ajpm.v10i2.3488

El-Adl, A. M., & Saad, M. A. E. (2019). Effect of a brain-based learning program on working memory and academic motivation among tenth grade Omanis students. *Online Submission*, *8*(1), 42-50.

El-Adl, A., & Alkharusi, H. (2020). Relationships between self-regulated learning strategies, learning motivation and mathematics achievement. *Cypriot Journal of Educational Sciences, 15*(1), 104-111. https://doi.org/10.18844/cjes.v15i1.4461

Engelbrecht, J., Llinares, S., & Borba, M. C. (2020). Transformation of the mathematics classroom with the internet. *ZDM, 52*(5), 825-841. https://doi.org/10.1007/s11858-020-01176-4

Farooq, S., Tatlah, I. A., & Butt, I. H. (2020). Role of Peer Tutoring on the Intrinsic Motivation of Student Teachers in Pakistan: An Experimental Investigation. *Pakistan Social Sciences Review, 4*(1), 381-388. https://doi.org/10.35484/pssr.2020(4-I)30

Ferreira, R. A., & Rodríguez, C. (2022). Effect of a science of learning course on beliefs in neuromyths and neuroscience literacy. *Brain Sciences, 12*(7), 1-14. https://doi.org/10.3390/brainsci12070811

Froiland, J. M., & Worrell, F. C. (2016). Intrinsic motivation, learning goals, engagement, and achievement in a diverse high school. *Psychology in the Schools, 53*(3), 321-336. https://doi.org/10.1002/pits.21901

Glaser, J. I., Benjamin, A. S., Farhoodi, R., & Kording, K. P. (2019). The roles of supervised machine learning in systems neuroscience. *Progress in Neurobiology, 175*, 126-137. https://doi.org/10.1016/j.pneurobio.2019.01.008

Hair Jr, J. F., Hult, G. T. M., Ringle, C. M., Sarstedt, M., Danks, N. P., & Ray, S. (2021). *Partial least squares structural equation modeling (PLS-SEM) using R: A workbook*. Springer. https://doi.org/10.1007/978-3-030-80519-7

Heyder, A., Weidinger, A. F., Cimpian, A., & Steinmayr, R. (2020). Teachers' belief that math requires innate ability predicts lower intrinsic motivation among low-achieving students. *Learning and Instruction, 65*, 1-39. https://doi.org/10.1016/j.learninstruc.2019.101220

Hofer, B. K. (2022). Teaching Educational Psychology to Undergraduates. Routledge. https://doi.org/10.4324/9781138609877-REE100-1

Hung, C. Y., Sun, J. C. Y., & Liu, J. Y. (2019). Effects of flipped classrooms integrated with MOOCs and game-based learning on the learning motivation and outcomes of students from different backgrounds. *Interactive Learning Environments, 27*(8), 1028-1046. https://doi.org/10.1080/10494820.2018.1481103

Immordino-Yang, M. H., Darling-Hammond, L., & Krone, C. R. (2019). Nurturing nature: How brain development is inherently social and emotional, and what this means for education. *Educational Psychologist*, *54*(3), 185-204. https://doi.org/10.1080/00461520.2019.1633924

Jailani, M. (2021). Developing Arabic Media Based on Brain-Based Learning: Improving Mufrodat in School. *Tadris: Jurnal Keguruan dan Ilmu Tarbiyah*, *6*(2), 349-361. https://doi.org/10.24042/tadris.v6i2.9921

Khan, R. M. G. B., Chachar, G. B., & Abro, I. A. (2020). Mathematics Achievement of Grade VIII Students Based on International Standardized Test (TIMSS) in an Urban Context of Sindh, Pakistan. *In 2020 3rd International Conference on Computing, Mathematics and Engineering Technologies (iCoMET),* 1-4. https://doi.org/10.1109/iCoMET48670.2020.9074110

Latoszek, B. B. V. (2020). Treatment effectiveness of Novafon local vibration voice therapy for dysphonia treatment. *Journal of Voice, 34*(1), 7-14. https://doi.org/10.1016/j.jvoice.2018.05.009

Li, P., & Lan, Y. J. (2022). Digital language learning (DLL): Insights from behavior, cognition, and the brain. *Bilingualism: Language and Cognition*, *25*(3), 361-378. https://doi.org/10.1017/S1366728921000353

Maass, K., Geiger, V., Ariza, M. R., & Goos, M. (2019). The role of mathematics in interdisciplinary STEM education. *ZDM, 51*(6), 869-884. https://doi.org/10.1007/s11858-019-01100-5

Mekarina, M., & Ningsih, Y. P. (2017). The effects of brain-based learning approach on motivation and students achievement in mathematics learning. *Journal of Physics: Conference Series, 895*(1), 1-6. https://doi.org/10.1088/1742-6596/895/1/012057

Mushtaq, T. (2021). A Critical Evaluation of PEELI Project Run by the Punjab Government of Pakistan and British Council. *International Journal of Academic Research in Business and Social Sciences, 11*(3), 1080-1098. https://doi.org/10.6007/IJARBSS/v11-i3/8656

Musu-Gillette, L. E., Wigfield, A., Harring, J. R., & Eccles, J. S. (2015). Trajectories of change in students' self-concepts of ability and values in math and college major choice. *Educational Research and Evaluation, 21*(4), 343-370. https://doi.org/10.1080/13803611.2015.1057161

Nematillayevna, J. M. (2021). Formation of methodological competencies of future mathematics teachers in the field of quality assessment of education. *Asian Journal of Multidimensional Research, 10*(6), 67-71. https://doi.org/10.5958/2278-4853.2021.00504.8

Oljayevna, O., & Shavkatovna, S. (2020). The Development of Logical Thinking of Primary School Students in Mathematics. *European Journal of Research and Reflection in Educational Sciences*, *8*(2), 235-239.

Palser, E. R., Lazerwitz, M., & Fotopoulou, A. (2022). Gender and geographical disparity in editorial boards of journals in psychology and neuroscience. *Nature Neuroscience*, *25*(3), 272-279. https://doi.org/10.1038/s41593-022-01012-w

Pascual, E. A. (2022). Getting the Answer Exactly Right: Dealing with Math Misconception. *Getting the Answer Exactly Right: Dealing with Math Misconception, 93*(1), 306-310. https://doi.org/10.47119/IJRP100931120222745

Pohan, A. M., Asmin, A., & Menanti, A. (2020). The effect of problem-based learning and learning motivation of Mathematical problem-solving skills of class 5 students at SDN 0407 Mondang. *Budapest International Research and Critics in Linguistics and Education (BirLE) Journal, 3*(1), 531-539. https://doi.org/10.33258/birle.v3i1.850

Ryan, R. M., & Deci, E. L. (2016). Facilitating and hindering motivation, learning, and well-being in schools: Research and observations from self-determination theory. *Handbook of Motivation at School*, 96-119. Routledge.

Saleh, S., & Mazlan, A. (2019). The effects of brain-based teaching with i-think maps and brain gym approach towards physics understanding. *Jurnal Pendidikan IPA Indonesia, 8*(1), 12-21. https://doi.org/10.15294/jpii.v8i1.16022

Sanetti, L. M. H., & Collier-Meek, M. A. (2019). Increasing implementation of science literacy to address the research-to-practice gap in school psychology. *Journal of School Psychology*, *76*, 33-47. https://doi.org/10.1016/j.jsp.2019.07.008

Solihatin, E., & Syahrial, Z. (2019). The effects of Brain-based learning and Project-based learning strategies on student group mathematics learning outcomes student visual learning styles. *Pedagogical Research*, *4*(4), em0047. https://doi.org/10.29333/pr/5949

Stojanović, J., Petkovic, D., Alarifi, I. M., Cao, Y., Denic, N., Ilic, J., ... & Milickovic, M. (2021). Application of distance learning in mathematics through adaptive neuro-fuzzy learning method. *Computers & Electrical Engineering, 93*, 1-12. https://doi.org/10.1016/j.compeleceng.2021.107270

Talpur, N., Kalwar, T., & Talpur, M. J. (2021). Computer-assisted Language Learning in Pakistani Context During COVID-19 Pandemic. REiLA: *Journal of Research and Innovation in Language, 3*(3), 210-225. https://doi.org/10.31849/reila.v3i3.6908

Tan, Y. S. M., & Amiel, J. J. (2022). Teachers learning to apply neuroscience to classroom instruction: a case of professional development in British Columbia. *Professional Development in Education, 48*(1), 70-87. https://doi.org/10.1080/19415257.2019.1689522

Tokac, U., Novak, E., & Thompson, C. G. (2019). Effects of game-based learning on students' mathematics achievement: A meta-analysis. *Journal of Computer Assisted Learning, 35*(3), 407-420. https://doi.org/10.1111/jcal.12347

Tortella, G. R., Seabra, A. B., Padrão, J., & Diaz-San Juan, R. (2021). Mindfulness and other simple neuroscience-based proposals to promote the learning performance and mental health of students during the COVID-19 pandemic. *Brain sciences*, *11*(5), 1-20. https://doi.org/10.3390/brainsci11050552

Triana, M., & Zubainur, C. M. (2019). Students' Mathematical Communication Ability through the Brain-Based Learning Approach Using Autograph. *Journal of Research and Advances in Mathematics Education, 4*(1), 1-10. https://doi.org/10.23917/jramathedu.v4i1.6972

Udjaja, Y., Guizot, V. S., & Chandra, N. (2018). Gamification for elementary mathematics learning in Indonesia. *International Journal of Electrical and Computer Engineering (IJECE), 8*(6), 3860-3865. https://doi.org/10.11591/ijece.v8i5.pp3860-3865

Weeden, K. A., Gelbgiser, D., & Morgan, S. L. (2020). Pipeline Dreams: Occupational Plans and gender differences in STEM major persistence and completion. *Sociology of Education*, *93*(4), 297-314. https://doi.org/10.1177/0038040720928484

Weidinger, A. F., Steinmayr, R., & Spinath, B. (2017). Math grades and intrinsic motivation in elementary school: A longitudinal investigation of their association. *British Journal of Educational Psychology, 87*(2), 187-204. https://doi.org/10.1111/bjep.12143

Wilcox, G., Morett, L. M., Hawes, Z., & Dommett, E. J. (2021). Why educational neuroscience needs educational and school psychology to effectively translate neuroscience to educational practice. *Frontiers in Psychology*, *11*, 618449. https://doi.org/10.3389/fpsyg.2020.618449

Wilkie, K. J., & Sullivan, P. (2018). Exploring intrinsic and extrinsic motivational aspects of middle school students' aspirations for their mathematics learning. *Educational Studies in Mathematics, 97*(3), 235-254. https://doi.org/10.1007/s10649-017-9795-y

Yu, R., & Singh, K. (2018). Teacher support, instructional practices, student motivation, and mathematics achievement in high school. *The Journal of Educational Research, 111*(1), 81-94. https://doi.org/10.1080/00220671.2016.1204260

Zakariya, Y. F., & Massimiliano, B. (2021). Development of Mathematics Motivation Scale: A Preliminary Exploratory Study with a Focus on Secondary School Students. *International Journal of Progressive Education, 17*(1), 314-324. https://doi.org/10.29329/ijpe.2021.329.20

iejee

International Electronic Journal of Elementary Education
September 2023, Volume 16, Issue 1, 123-141

An Investigation of Middle School Students' Spatial Reasoning Skills

Gamze Kurt[a,*], Fatih Önel[b], Özge Çakıoğlu[c]

Received : 11 April 2023
Revised : 19 September 2023
Accepted : 29 September 2023
DOI : 10.26822/iejee.2023.319

[a,*] **Corresponding Author:** Gamze Kurt, Mersin University, Faculty of Education, Mersin, Türkiye.
E-mail: gamzekurt@mersin.edu.tr
ORCID: https://orcid.org/0000-0002-4976-5069

[b] Fatih Önel, Ministry of National Education (MoNE), Mersin, Türkiye.
E-mail: fatihonel@gmail.com
ORCID: https://orcid.org/0000-0002-0560-8141

[c] Özge Çakıoğlu, Ministry of National Education (MoNE), Mersin, Türkiye.
E-mail: cakiogluozge@gmail.com
ORCID: https://orcid.org/0000-0002-0355-8371

Abstract

In recent years, there has been a growing interest in spatial reasoning as a component of mathematics education, with many countries incorporating it into their mathematics curriculum. The study of spatial reasoning in the learning areas of geometry presents an essential opportunity for improvement of students within the realm of mathematics education. The purpose of this survey research is to examine middle school students spatial reasoning skills. A study was conducted to analyze the spatial reasoning skills of 947 middle school students. The Spatial Reasoning Test was utilized to assess the sub-components of spatial visualization, spatial orientation, and mental rotation. Based on the results, there was no significant difference between male and female students in relation to their overall test scores. However, a statistically significant difference was observed when analyzing the scores of the sub-components and grade levels. Anticipated outcomes of the investigation are expected to provide support and guidance for the preparation of educational tasks and instruction aimed at enhancing students' spatial reasoning abilities.

Keywords:

Spatial Reasoning, Middle School Students, Mental Rotation, Spatial Orientation, Spatial Visualization

Introduction

The concept of spatial reasoning is one that is approached from a variety of disciplinary perspectives. Despite being explored in disciplines other than mathematics, spatial reasoning has been an integral part of mathematics education research and the mathematics curriculum for decades. There have been various definitions of spatial reasoning, including spatial ability, spatial perception, spatial reasoning, three-dimensional thinking, and spatial perception (Clements & Battista, 1992; NCTM, 2000; Olkun, 2003). In the literature, the terms spatial reasoning, spatial skills, and spatial ability are frequently used interchangeably. Since the concept of ability is used to express inherent ability, the term spatial reasoning skill was chosen for this research. In recent years, spatial reasoning skills, which are considered part of geometry education, have become increasingly essential, and it has been discovered that geometry teaching plays a crucial role in the development of spatial reasoning skills (Clements & Sarama, 2011). The inadequacy of teachers' geometry content knowledge (Clements & Sarama, 2011) and the reduction of middle

school geometry instruction to the properties and relationships of two- and three-dimensional shapes (Sinclair & Bruce, 2015) may also impact the teaching of spatial reasoning skills.

When examining the teaching of spatial reasoning concepts in Turkey from the past to the present, the concept of symmetry, which is one of the sub-topics, was first included in the 1968 mathematics curriculum, and in the 2009 mathematics curriculum, it was intended to be taught by associating it with geometric shapes, and it was explicitly stated that it was taught within the scope of the transformation geometry sub-learning area (Memişoğlu & Tapan-Broutin, 2018). In the mathematics curriculum, spatial reasoning skills are developed through the sub-learning areas 'spatial relations' at the elementary school level, 'transformation geometry' at the middle school level, and 'fundamental transformations at the analytical plane' at the high school level. The aforementioned sub-learning areas include the instruction of concepts such as expressions indicating location and direction, symmetry, line of symmetry, mirror symmetry, reflection, translation, image, center of rotation, angle of rotation, axis of symmetry, and center of symmetry (MoNE, 2018). In addition to these, spatial reasoning skills include drawing two-dimensional views of three-dimensional objects from different directions, creating structures based on drawings of their views from different directions, recognizing three-dimensional geometric objects and drawing their expansion, and producing images of points, line segments, and other shapes in the plane as a result of translation, reflection, and rotation (MoNE, 2018).

By acquiring geometric thinking skills, it is planned that students will be able to establish relationships between spatial reasoning skills such as critical thinking, creative thinking, and multi-dimensional thinking and other areas of mathematics (MoNE, 2018). The geometry learning area is an essential component of the mathematics curriculum and provides students with a significant opportunity to develop their spatial abilities. Numerous studies demonstrate the link between spatial reasoning and achievement in mathematics, engineering, science, and technology (STEM) (e.g. Fowler et al., 2022; Mix & Cheng, 2012; Shea et al., 2001; Wai et al., 2009).

The curriculum emphasizes the importance of developing spatial visualization and interpretation skills in all students in order for them to succeed in the mathematics course (MoNE, 2018; NCTM, 2000). Although spatial reasoning is included in our curriculum, many students struggle with spatial reasoning questions (Kabakçı & Demirkapı, 2016). The fact that concepts related to spatial reasoning have been studied for a shorter period of time than other concepts in the field of learning geometry calls for some modifications to the curriculum. Incorporating more tools, such as dynamic geometry software, into geometry instruction in recent years demonstrates the need for new goals in geometry education currently. It is also mentioned that success in international applications such as Programme for International Student Assessment (PISA) scores are correlated with spatial reasoning test scores, and that this correlation may be an undervalued strategy for improving performance on these examinations (Sorby & Panther, 2020). Lowrie and Logan (2018) recommend associating geometry concepts with spatial reasoning more in countries that excel at international student success assessments such as Trends in International Mathematics and Science Study (TIMMS) or PISA. From this perspective, the emphasis placed on geometry instruction will also contribute to the growth of students' spatial reasoning abilities. In order to cultivate students' spatial reasoning skills, it is first necessary to assess their current level of spatial reasoning ability so that teaching innovations can be planned.

Geometry teaching allows students to spatially reason about geometric concepts (Clements & Battista, 1992). Geometry enables students to understand, model, and manipulate the structure of objects, shapes, and space. Therefore, teaching geometry should be viewed as a significant opportunity to foster spatial reasoning in students. Rotating, reflecting, and situating two-dimensional shapes activates students' spatial visualization skills, and spatial visualization can be enhanced by associating this skill with concepts like the area of shapes and their positions on the analytical plane (Lowrie & Logan, 2018). As an example of the relationship between spatial reasoning skills and other learning areas, it has been observed that students with strong mental rotation skills perform better on algebraic thinking problems (Cheng & Mix, 2014). In order to increase achievement in other areas of mathematics learning, it is evident that spatial reasoning abilities must be developed. At both the primary and secondary school levels, it is important to develop geometric reasoning skills holistically by enhancing students' knowledge of geometry concepts, as opposed to simply developing spatial reasoning skills; therefore, appropriate learning environments should be associated with geometry concepts for the development of spatial reasoning (Fujita et al., 2020). The development of spatial reasoning skills should begin as early as possible (Casey et al., 2008), as good spatial reasoning skills lead to good geometry achievement and therefore mathematical success (Burte et al., 2017; Mulligan, 2015). In order to develop spatial reasoning, it is necessary to assess the current level of spatial reasoning ability among students. In addition, the level of students' spatial reasoning skills must be evaluated in order to construct an effective geometry education.

Although there is no consensus regarding the subcomponents it consists of as a complex concept and the definitions given, three fundamental subcomponents of spatial reasoning skill can be emphasized based on the concepts in primary and middle school mathematics curriculum: mental rotation, spatial orientation, and spatial visualization (Ramful et al., 2017). The purpose of this survey study is to determine the spatial reasoning skills of middle school students and provide findings to support the necessary preparation for the development of spatial reasoning. In addition, it was intended to determine which subcomponent of spatial reasoning may be responsible for difficulties in developing spatial reasoning skills, as well as whether these subcomponents vary by gender and grade level. To achieve these objectives, the following research questions will be answered:

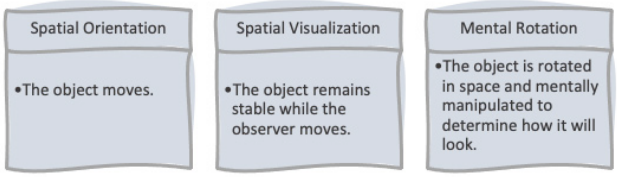
1. How well do middle school students incorporate spatial reasoning?
2. Do female and male middle school students score differently on spatial reasoning?
3. Are the spatial reasoning skill scores of middle school students different in terms of the grade level they attend?
4. Do female and male middle school students score differently on spatial visualization, spatial orientation, and mental rotation subcomponents of spatial reasoning?

Theoretical Background

Spatial reasoning can be defined as the capacity to create or manipulate images or shapes. This ability encompasses the understanding and manipulation of three-dimensional objects and their positions and relationships in space. Consequently, the creation or orientation of shapes or images is only a portion of the process. Success in mathematics, engineering, science, and technology requires the development of spatial reasoning. The spatial reasoning ability, which develops simultaneously with the geometric reasoning ability, is defined as "the ability to recognize, produce, analyze, operate on, and reflect on spatial objects, images, relations, movements, and transformations" (Battista et al., 2018). Following is a brief description of additional accepted definitions from the literature: While Lohman (1996) defines spatial reasoning as the ability to create a visual image, maintain a given shape, and transform it into another shape, McGee (1979) defines it as the ability to visualize three-dimensional objects and their movements. Van De Walle et al. (2012) defined it as the ability to perceive objects from various perspectives, comprehending the relationship between two-dimensional and three-dimensional structures, and imagining the open and closed forms of objects. Spatial reasoning was defined by Turğut (2007) as the capacity to mentally manipulate shapes and generate explanations by rearranging them. For instance, spatial reasoning is used when mentally animating, rotating, or relocating an object or shape. Finding the expansion of a cube, locating a given shape or location on a map, locating the reflection of a given point, and determining its symmetry all require spatial reasoning. Due to its varied and complex definitions, spatial reasoning is comprised of numerous subcomponents. These include spatial visualization, spatial orientation, mental rotation, spatial perception, spatial relations, mental rotation, and spatial cognition (Kayhan, 2005; Lohman, 1979; McGee, 1979; Okagaki & Frensch, 1994; Turğut, 2007;).

Although this study does not intend to investigate the subcomponents of spatial reasoning, it will adhere to a three-component theoretical framework: spatial orientation, spatial visualization, and mental rotation (Ramful et al., 2017), which are shown in Figure 1 below. Spatial orientation is the ability to envision how a particular object or group of objects will appear from various perspectives (McGee, 1979; Lohman, 1979; Lowrie & Logan, 2018). It requires the individual to mentally reposition himself in order to make sense of and interpret visual representations of objects, such as when using a map (Pietropaolo & Crusio, 2012). Second, spatial visualization is the ability to perform complex mental transformations, such as folding paper in the mind, and to imagine what an object or shape will become after being spatially transformed or changed into a different shape (Clements, 1998; Lohman, 1979; McGee, 1979; Ramful et al., 2017). The individual must be able to visualize the object's final form after multiple transformations. Lastly, mental rotation is a cognitive action that comes into play in situations such as imagining how a two- or three-dimensional object will appear when rotated from a particular angle (Lowrie & Logan, 2018; Okagaki & Frensch, 1994). This can occur with two distinct tasks: perspective tasks, in which the individual considers how the object will appear from a different angle, and comparison tasks, which deal with changes in the object itself as opposed to the individual's position or perspective (Fowler et al., 2022; Guillot et al., 2012).

Figure 1.
Subcomponents of spatial reasoning



This diversity in the definition and components of the concept of spatial reasoning has also led to differences in the tests used to ascertain the level of this skill (Yurt & Sünbül, 2012). In a number of studies conducted in the area of spatial reasoning, the relationships between

this ability and various variables were investigated. The effect of spatial reasoning skills on mathematics achievement has been studied and debated for quite a while. There is a significant relationship between spatial reasoning skills and mathematics achievement, as shown by studies (Battista, 1990; Cheng & Mix, 2014; Kayhan, 2005; Turğut & Yılmaz, 2012).

Baki et al. (2011) concluded that teaching spatial visualization skills using virtual manipulatives and dynamic geometry software is more effective than the traditional method. Similarly, it has been demonstrated that teachers should utilize dynamic geometry software to enhance their students' spatial reasoning abilities (Güven & Kösa, 2008). In addition, research indicates that the use of a technologically prepared STEM program has a positive effect on the spatial abilities of seventh graders (Fowler et al., 2022). Additionally, it has been noted that incorporating real-world examples and computer programs in courses enhances students' spatial skills (Yıldız & Tüzün, 2011; Yolcu, 2008). According to Casey et al. (2008), there is a close relationship between spatial reasoning skills and success in geometry and mathematics, and mental rotation, spatial visualization, and building construction processes are the basis of this relationship. In their study of eighth grade students, Yıldırım Gül and Karataş (2015) concluded that there is a positive and significant relationship between students' geometry understanding levels, geometry achievement levels, and spatial reasoning abilities. Cheng and Mix (2014) reached the conclusion that activities aimed at fostering mental rotation skill, which is a subcomponent of spatial reasoning skill, improve the performance of students in early childhood, particularly in missing term problems. In addition, it has been determined that it is essential to design and evaluate spatial reasoning programs in order to enhance the mathematics learning of primary school students as they progress through the grade levels (Woolcott et al., 2022).

There are also studies that attempt to determine the relationship between spatial reasoning ability and other variables, including gender, pre-school education status, grade level, early involvement with mechanical games, teaching method, use of three-dimensional virtual environments, and spatial reasoning ability levels of teachers. These studies demonstrated that such variables influence the spatial abilities of students (Bartlett & Camba, 2023; Ben-Chaim et al., 1988; Sorby, 1999; Turğut, 2007; Yurt & Sünbül, 2012). Studies show that students' spatial reasoning skills develop as their grade level progresses (Akkaya Yılmaz, 2022; Turğut & Yılmaz, 2012). When the pre-school education status was examined, it was determined that the pre-school education of the individual had a positive effect on spatial reasoning skills and that mechanical games included in early childhood education had a positive effect on spatial reasoning skills (Akkaya Yılmaz, 2022; Turğut, 2007). Ben-Chaim et al. (1988) suggest that the teaching method influences spatial reasoning ability positively and that the seventh grade is the optimal time for spatial reasoning-based education. According to Yurt and Sünbül (2012), activities involving concrete objects have a positive influence on students' spatial reasoning abilities.

The majority of research examining the effect of the gender variable on spatial reasoning abilities demonstrates a positive difference in favor of males (Ben-Chaim et al., 1988; Sorby, 1999). Studies on the measurement of spatial reasoning with its subcomponents (Turğut et al., 2017; Voyer & Doyle, 2010) are also included in the literature, and in the mental rotation subcomponent, males perform better than females. The examination of spatial reasoning and its subcomponents has yielded varying results depending on the gender variable in several studies (Kaya, 2019; Seng & Chan, 2000;). In contrast to these studies, others have shown that there is no difference between male and female pupils in terms of mental rotation skill, which is a subcomponent of spatial skills (e.g. İrioğlu & Ertekin, 2012; Yıldız, 2009). However, in a recent study, Bartlett and Camba (2023) obtained results that contradict the studies that found results in favor of girls or boys on spatial reasoning skills. They assert that the previous findings are due to the masculine structure of the tests or the development of gender roles and spatial reasoning skills together. Moè (2009) concluded that students' perspectives on spatial reasoning influence the results of the spatial rotation test, a subcomponent of spatial reasoning. Moreover, Ramful et al. (2017) and Turğut (2007) have demonstrated that distinct sub-components of spatial reasoning are interrelated.

Studies indicate that students' spatial reasoning skills are inadequate (Uygan & Turğut, 2012) and that personal and environmental factors are responsible for this deficiency (Turğut & Yılmaz, 2012). In addition to the relational or survey studies summarized in this section, there are studies examining the effects of numerous teaching methods designed to improve spatial reasoning skills on spatial reasoning (Atasoy et al., 2019; Casey et al., 2008; Gün & Atasoy, 2017; Uygan, 2011; Yıldız & Tüzün, 2011). Yıldız (2009) asserted that the use of a three-dimensional virtual environment improved spatial visualization ability, and Turhan (2010) claimed that computer-aided perspective drawings had a positive effect on students' attitudes toward mathematics, technology, and geometry. Skill in spatial reasoning can be acquired and enhanced through specially designed interventions; therefore, it is dependent on experience (Lowrie et al., 2019; Uttal et al., 2013). For instance, concrete models and computer applications can improve the spatial reasoning skills

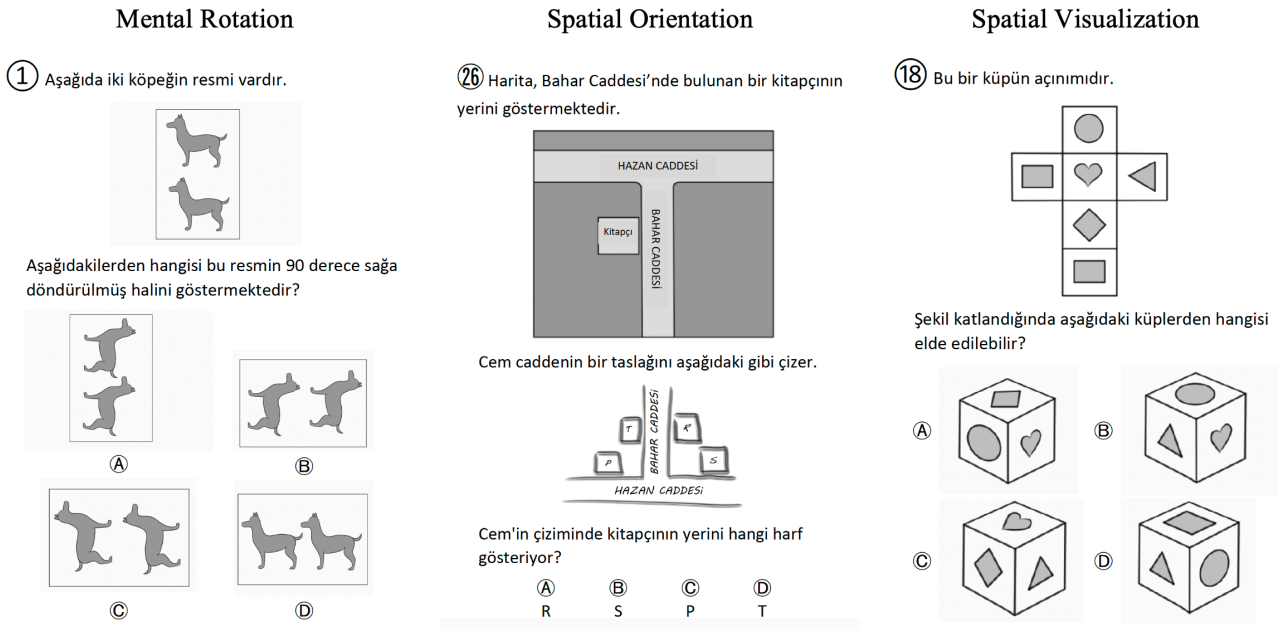
of sixth-grade students (Yolcu, 2008), and the use of concrete materials and three-dimensional media has a significant influence on mental rotation and spatial visualization skills (Yıldız & Tüzün, 2011). Alternatively, it has been established that by constructing dynamic items with augmented reality, students can observe objects from various angles, resulting in active learning (Anggraini et al., 2020). High-level cognitive skills are required for spatial reasoning, so related concepts can be embodied in two- or three-dimensional visuals or objects (Baki, 2000). The relationship between spatial reasoning ability and mathematics achievement is consistent and strengthens over time (Resnick et al., 2019; Resnick et al., 2020).

Method

In this study, one of the quantitative research approaches, survey research, was employed. Survey research is a research method that tries to describe the characteristics of a certain group in detail through variables (Gravetter & Forzano, 2018). This survey explores the spatial reasoning abilities of middle school students. The spatial reasoning skills of middle school students in the 5th, 6th, 7th, and 8th grades are examined in detail using the three subcomponents of spatial reasoning, spatial visualization, spatial orientation, and mental rotation, which are explained in detail in the above theoretical framework. All students from a middle school compose the sample. Students differ in age from 10 to 13 years old. In accordance with the research permission granted by the Provincial Directorate of National Education, 947 middle school students were recruited and the test was administered simultaneously in the entire school on the previously scheduled date. Before the allocated practice time, clear test information was provided to the teachers in the teachers' room, and they were asked to check the students' answer sheets to prevent marking errors. The application was completed within an hour of class time, and the test forms brought by the teachers were collected and organized during the subsequent break.

This study's data collection tool is the spatial reasoning test, which describes the spatial reasoning skills of middle school students with the three subcomponents mentioned. The data collection tool includes a total of 30 items, consisting of ten items designed to measure three separate components. Ramful et al. (2017) designed the Spatial Reasoning Test (SRT) and obtained an internal reliability value of .845. In addition, the SRT correlates with all three components as follows: mental rotation (.71), spatial orientation (.41), and spatial visualization (.66). Prior to the application, SRT was translated from its original language, English, into Turkish, and in certain items, attention was given to include Turkish names of places and people so that the context would be compatible with Turkish. Since SRT is intended for middle school students, attention has been paid to include clear and simple expressions in both the original and the translation. The Turkish SRT was presented to an expert in mathematics education research, and his professional opinion was requested. After the corrections were made based on their feedback, the test was prepared as 32 pages in total, including the pages where coding and demographic information were requested, with one item per page, identical to the original test. Figure 2 provides examples of items, one for each of the three subcomponents:

Figure 2.
Sample questions for subcomponents of mental rotation spatial orientation, and spatial visualization



The data obtained by the implementation of SRT was analyzed in detail with the Jamovi program, which is a R (R Core Team, 2022) based application (The Jamovi Project, 2023). SnowIRT, one of Jamovi's modules, was employed for Rasch analysis (Seol, 2023). Various data representations, including descriptive statistics and histograms or box plots, were also constructed and analyzed for the participants' SRT total scores, gender, grade level, and spatial reasoning subcomponents.

The data obtained from SRT were modelled using the unidimensional one-parameter logistic model known as the Rasch model with the marginal maximum likelihood method (de Ayala, 2009). There are some assumptions of conducting Rasch analysis: unidimensionality, local independence and monotonicity. Yen's Q3 statistic was measured to check assumption of local independence (Yen, 1984). When the residual correlation matrix was also examined, the assumption of local independence was achieved because the residual values of the items were below .20. In other words, none of the items in test are related to each other.

Findings

The findings obtained in this section were analyzed in accordance with the study's research questions. Table 1 displays the minimum and maximum values, arithmetic mean, median, standard deviation, skewness, and kurtosis values of the total scores of 947 students who were administered the scale.

Table 1.
Summary of total scores

N	Min	Max	Mean	Median	SD	SE	Skewness	Kurtosis
947	1.00	28.0	13.4	13.0	4.91	0.160	0.488	2.74

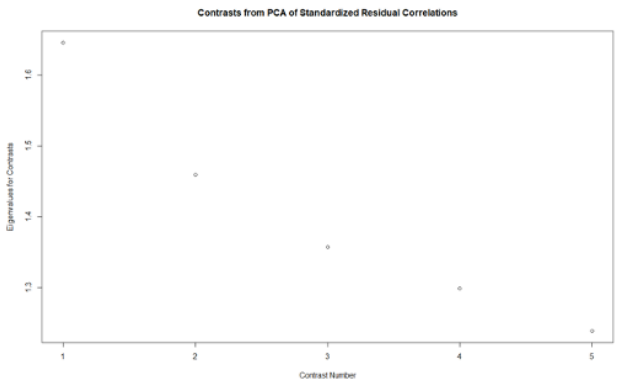
The dichotomous Rasch model was used to conduct item analyses of the SRT, which is used to understand how items and participants operate simultaneously. As shown in Table 2, the items were coded with 1 (for correct results) and 0 (for incorrect results) and dichotomous Rasch model was used to analyze these dichotomous item data. According to Table 2, it can be said that the assumption of local independence was not violated (Christensen et al., 2017).

Table 2.
Yen's Q3 statistic based on pearson correlation

Mean	Max	Min	Max_abs	Min_abs	Q3
-0.0329	0.136	-0.135	0.136	1.08e-4	0.169

We conducted principal components analysis of standardized residual correlations (PCA of residuals in R to check the assumption of unidimensionality in dichotomous Rasch model. As shown in the Figure 3 all of the contrasts are smaller than 2.00 (Linacre, 2016). Then, it can be said that the unidimensionality assumption was not violated.

Figure 3.
Contrasts from PCA of standardized residual correlations



The fact that the infit and outfit values are close to one indicates that the items are well-prepared, as shown in Table 3. These values satisfy the criteria for item fit and demonstrate that each item contributes to the comprehension of spatial reasoning ability. In addition, logit values indicating the item's difficulty level displayed for each item. Positive values emphasize more difficult items, while negative values highlight easier ones.

Table 3.
Item Statistics

	Proportion	Measure (logit)	SE Measure	Infit	Outfit
ZD1	0.599	-0.753	0.0715	1.032	1.037
ZD4	0.617	-0.840	0.0720	0.947	0.926
ZD7	0.461	-0.103	0.0708	0.944	0.930
ZD10	0.362	0.382	0.0734	0.974	0.957
ZD13	0.472	-0.153	0.0706	0.950	0.937
ZD16	0.189	1.421	0.0889	0.987	0.951
ZD19	0.214	1.240	0.0851	0.915	0.903
ZD22	0.299	0.720	0.0769	0.888	0.884
ZD25	0.287	0.786	0.0777	1.044	1.104
ZD28	0.231	1.127	0.0830	1.023	1.149
UY2	0.859	-2.332	0.0972	0.931	0.920
UY5	0.799	-1.870	0.0855	0.902	0.817
UY8	0.800	-1.878	0.0857	0.968	0.961
UY11	0.501	-0.287	0.0705	0.963	0.950
UY14	0.688	-1.206	0.0750	0.882	0.826
UY17	0.641	-0.961	0.0728	0.907	0.873
UY20	0.749	-1.545	0.0795	1.091	1.311
UY23	0.570	-0.616	0.0709	0.856	0.815
UY26	0.388	0.255	0.0724	0.978	0.973
UY29	0.388	0.255	0.0724	1.090	1.135
UG3	0.695	-1.240	0.0754	1.005	1.042
UG6	0.316	0.627	0.0758	0.994	1.052
UG9	0.321	0.598	0.0755	1.055	1.068
UG12	0.326	0.570	0.0752	0.921	0.911
UG15	0.399	0.197	0.0721	1.101	1.108
UG18	0.322	0.593	0.0754	1.167	1.257
UG21	0.175	1.527	0.0914	1.182	1.508
UG24	0.268	0.897	0.0793	1.075	1.164
UG27	0.184	1.461	0.0898	1.015	1.141
UG30	0.231	1.127	0.0830	1.085	1.361

The person-item map and graphs displaying the infit-outfit values of the items were depicted in Figure 4, below.

The WrightMap depicted in Figure 5 was created based on the findings obtained, and Figure 5 depicts how the data were distributed based on the difficulty levels of the items. The person-item map, named the WrightMap in honor of Rasch measurement advocate Ben Wright, displays individuals (in terms of their abilities) and items (in terms of their difficulty) along a common (usually vertical) axis marked by a scale (Callingham & Bond, 2006). On the WrightMap logit scale, individuals' abilities and item difficulties are represented by estimates. Individuals in the same position on the logit scale (for instance, assuming item difficulty is 0.6) are aligned to form a long bar at the level of 0.6 on the scale, and according to the Rasch model, each individual (individuals shown in the bar) will have a 50% chance of responding to the relevant item correctly (Callingham & Bond, 2006). This map reveals that the second item (UY2) for the spatial orientation component is the least difficult, while the twenty-first item (UG21) for the spatial visualization component is the most difficult.

Table 4 provides the expected score curves for each item derived by applying the Rasch model. The expected score curves of the items reveal the progression of the item statistics in Table 3, as well as the difficulty levels of the items. Table 4 also shows the monotonicity of Rasch analysis which was achieved.

Figure 4.
Person-item map and Infit-Outfit values for the items

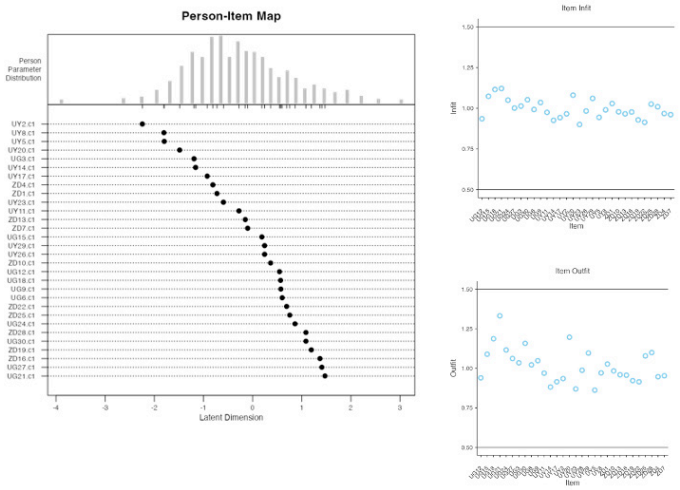


Figure 5.
WrightMap

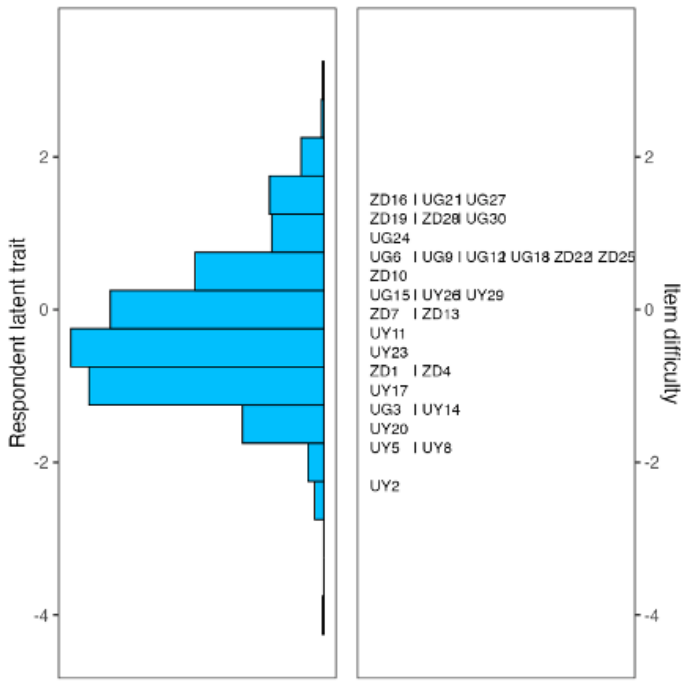
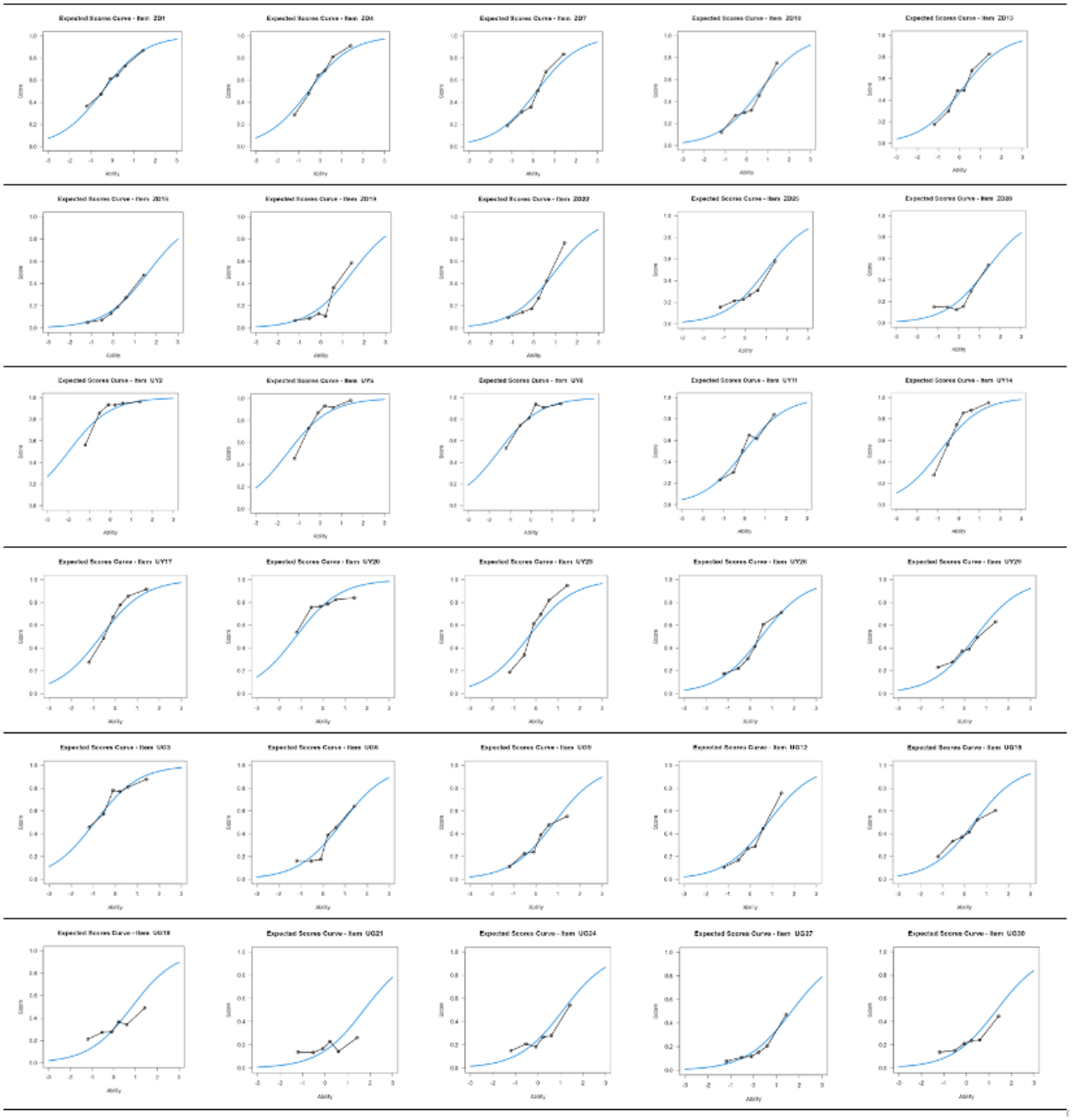


Table 4.
Expected score curve for all items



A. Findings regarding Reliability

The results of the test conducted to ascertain the reliability of SRT scores are provided in Table 5 below. The obtained Cronbach α value (α = .770) shows that the test results are reliable.

Since it is very close to the original reliability values (α =.849) of Ramful et al. (2017), SRT can be considered statistically reliable. Table 6 provides the results of the analysis demonstrating the contribution of each item to the reliability index.

B. Confirmatory Factor Analysis

The results of the confirmatory factor analysis conducted to determine whether the components specified in the original test (mental rotation, spatial orientation, and spatial visualization) were also maintained in the Turkish version of the SRT are presented in Table 7 as a validity indicator.

The fact that the p values for each item in Table 7 are less than α =.05 indicates that the described factors have been preserved and are consistent with the original test. According to Ramful et al. (2017), it has been confirmed that the SRT is an effective measurement instrument for all three components of spatial reasoning ability. Ramful et al. (2017) demonstrated that the 30-item SRT can assess mental rotation, spatial orientation, and spatial visualization independently. This part was conducted according

Table 5.
Reliability Statistics

	Mean	SD	Cronbach's α	McDonald's ω
Scale	0.445	0.164	0.770	0.771

Table 6.
Item Reliability Statistics

			If item dropped	
	Mean	SD	Cronbach's α	McDonald's ω
ZD1	0.599	0.490	0.766	0.766
ZD4	0.617	0.486	0.761	0.762
ZD7	0.461	0.499	0.759	0.760
ZD10	0.362	0.481	0.761	0.762
ZD13	0.472	0.499	0.760	0.761
ZD16	0.189	0.392	0.763	0.764
ZD19	0.214	0.411	0.760	0.760
ZD22	0.299	0.458	0.757	0.757
ZD25	0.287	0.453	0.766	0.767
ZD28	0.231	0.422	0.766	0.767
UY2	0.859	0.349	0.765	0.766
UY5	0.799	0.401	0.762	0.762
UY8	0.800	0.400	0.765	0.766
UY11	0.501	0.500	0.761	0.762
UY14	0.688	0.463	0.758	0.759
UY17	0.641	0.480	0.759	0.760
UY20	0.749	0.434	0.772	0.773
UY23	0.570	0.495	0.754	0.756
UY26	0.388	0.487	0.761	0.762
UY29	0.388	0.487	0.768	0.769
UG3	0.695	0.461	0.766	0.767
UG6	0.316	0.465	0.763	0.764
UG9	0.321	0.467	0.766	0.767
UG12	0.326	0.469	0.758	0.759
UG15	0.399	0.490	0.769	0.769
UG18	0.322	0.468	0.773	0.773
UG21	0.175	0.380	0.774	0.776
UG24	0.268	0.443	0.768	0.769
UG27	0.184	0.387	0.765	0.766
UG30	0.231	0.422	0.769	0.770

Table 7.
Factor Loadings

				%95 Confidence interval			
Factor	Indicator	Estimation	SE	Lower	Upper	Z	p
Mental Rotation	ZD1	1.000 [a]					
	ZD4	1.351	0.216	0.9280	1.773	6.26	<.001
	ZD7	1.622	0.245	1.1411	2.102	6.61	<.001
	ZD10	1.405	0.220	0.9738	1.837	6.38	<.001
	ZD13	1.467	0.229	1.0173	1.916	6.39	<.001
	ZD16	1.064	0.171	0.7285	1.400	6.21	<.001
	ZD19	1.362	0.205	0.9608	1.763	6.65	<.001
	ZD22	1.640	0.241	1.1676	2.113	6.80	<.001
	ZD25	0.933	0.171	0.5985	1.267	5.47	<.001
	ZD28	0.920	0.163	0.6001	1.240	5.63	<.001
Spatial Orientation	UY2	1.000 [a]					
	UY5	1.418	0.173	1.0790	1.757	8.20	<.001
	UY8	0.953	0.144	0.6700	1.236	6.60	<.001
	UY11	1.632	0.207	1.2269	2.037	7.90	<.001
	UY14	1.827	0.213	1.4102	2.245	8.59	<.001
	UY17	1.799	0.214	1.3796	2.218	8.41	<.001
	UY20	0.544	0.135	0.2787	0.810	4.02	<.001
	UY23	1.983	0.230	1.5332	2.433	8.64	<.001
	UY26	1.416	0.191	1.0425	1.789	7.43	<.001
	UY29	0.841	0.161	0.5263	1.156	5.23	<.001
Spatial Visualization	UG3	1.000 [a]					
	UG6	1.293	0.212	0.8776	1.709	6.10	<.001
	UG9	0.941	0.181	0.5874	1.296	5.21	<.001
	UG12	1.783	0.264	1.2661	2.301	6.76	<.001
	UG15	0.878	0.181	0.5238	1.232	4.86	<.001
	UG18	0.448	0.148	0.1584	0.737	3.03	0.002
	UG21	0.153	0.113	-0.0675	0.374	1.36	0.174
	UG24	0.825	0.166	0.4996	1.150	4.97	<.001
	UG27	0.887	0.159	0.5750	1.198	5.58	<.001
	UG30	0.681	0.150	0.3871	0.975	4.54	<.001

[a] fixed parameter

to the steps of confirmatory factor analysis, and the standard factor loadings that each item contributed to each factor were reported in the model data fit report. The obtained correlation values for mental rotation scores (MRS), spatial orientation scores (SOS), and spatial visualization scores (SVS) were also shown in Table 8 below.

Table 8.

Correlation Matrix

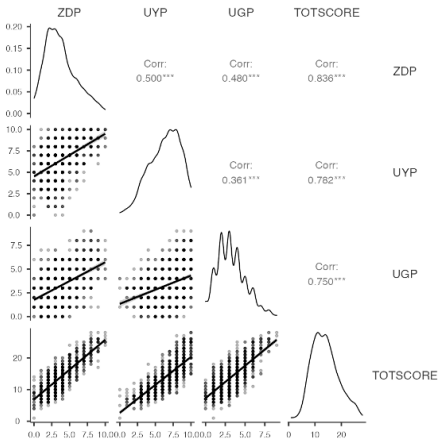
		MRS		SOS		SVS		TOTSCORE
MRS	Pearson's r	—						
	p-value	—						
	95% CI Upper	—						
	95% CI Lower	—						
SOS	Pearson's r	0.500	***	—				
	p-value	<.001		—				
	95% CI Upper	0.547		—				
	95% CI Lower	0.451		—				
SVS	Pearson's r	0.480	***	0.361	***	—		
	p-value	<.001		<.001		—		
	95% CI Upper	0.528		0.415		—		
	95% CI Lower	0.430		0.304		—		
TOTSCORE	Pearson's r	0.836	***	0.782	***	0.750	***	—
	p-value	<.001		<.001		<.001		—
	95% CI Upper	0.854		0.805		0.776		—
	95% CI Lower	0.816		0.755		0.720		—

Not. * p <.05, ** p <.01, *** p <.001

There is a statistically significant relationship between MRS, SOS, SVS, and total scores, as shown in Table 8. The correlation values range from 0.361 to 0.836. As with the original spatial reasoning skill test, these values are statistically significant, demonstrating the construct validity of the data collection tool (Ramful et al., 2017). Figure 6 depicts the correlations between MRS, SOS, SVS, and total scores (TOTSCORE), in addition to the table presented previously.

Figure 6.

Correlation among the variables of MRS, SVS, SOS and TOTSCORE



The results of the model fit analysis of SRT were presented in Table 9 below.

Table 9.

Ki-squared Test for exact fit

χ^2	df	p
559	402	<.001

Since the p value in Table 9 is less than.05, it is evident that the model defined with the three mentioned components has an acceptable level of fit. According to this model, χ^2 (chisquare) / degrees of freedom (df) = 559 / 402 equals 1.39. The fact that the value is less than 3 indicates that the model's goodness of fit is acceptable, but since it is insufficient, Table 10 provides additional goodness-of-fit measures.

Table 10.

Fit measures

			RMSEA 90% CI	
CFI	TLI	RMSEA	Lower	Upper
0.934	0.928	0.0203	0.0161	0.0242

Based on the results presented in Table 10, it can be concluded that the test is highly compatible with the specified model. CFI, TLI, and RMSEA values indicate that there is no problem with the model's fit. For example, as an item fit measure more appropriate for large samples, RMSEA (root mean square error of approximation) values of .06 or below suggests a strong level of model fit (Tabachnick & Fidell, 2019).

C. Is there a difference in total scores between girls and boys?

Descriptive statistics for the research question "Is there a difference between the total scores of girls and boys?" were presented in Table 11 and the corresponding histogram was depicted in Figure 7.

Table 11.

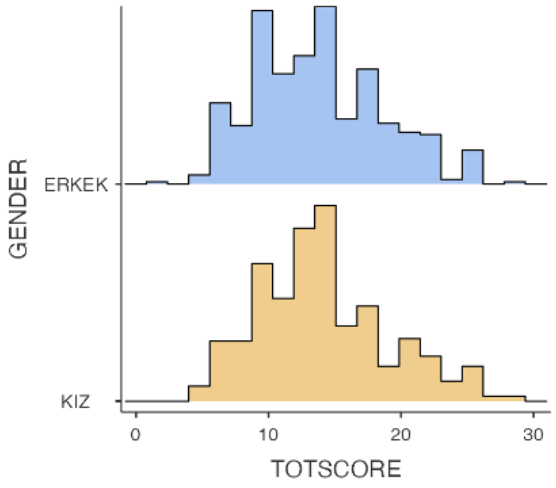
Descriptive Statistics

	Gender	N	Missing	Mean	Median	SD	Minimum	Maximum
TOTSCORE	MALE (ERKEK)	499	0	13.8	13	4.86	1	28
	FEMALE (KIZ)	448	0	14.1	14.0	4.97	4	28

Histograms of the distribution of the total scores obtained by male and female students were shown in Figure 7 below.

Figure 7.

Distribution of the total scores of females and males (histogram)



The Kolmogorov-Smirnov test of normality was used to test the normality assumption in order to determine whether the groups of male and female students exhibited a normal distribution, and the Levene test was used to test the homogeneity of variances; the results obtained were presented in Tables 12 and Table 13 below.

The Levene test result for the homogeneity of variances was shown in Table 13 below. The obtained p value indicates that the variances were not distributed homogeneously.

According to the results of these tests, the normality assumption of the parametric tests could not be satisfied, so the Mann-Whitney U test was employed to compare the groups. The test results were displayed in Table 14 below.

The null hypothesis that "there is no significant difference between girls and boys in terms of total scores" could not be rejected based on the data presented in the table above. In other words, there is no statistically significant difference between girls' and boys' total scores.

D. Is there a difference between the grade levels in terms of total scores?

Table 15 provides descriptive statistics in order to determine whether there is a difference between grade levels, which is the next question of the study. Examining Table 15 reveals that the number of students in each grade level in the sample was evenly distributed, and that the mean scores range from 12.2 to 15.2.

Table 12.
Normality Tests

		Statistics	p
TOTSCORE	Shapiro-Wilk	0.975	<.001
	Kolmogorov-Smirnov	0.0769	<.001
	Anderson-Darling	6.73	<.001

Table 13.
Homogeneity of Variances Test

		F	df	df2	p
TOTSCORE	Levene	0.0601	1	945	0.806
	Variance ratio	0.957	498	447	0.629

Table 14.
Results of Mann-Whitney U Test

		Statistics	df	p	Mean difference	SD difference		Effect size
TOTSCORE	Student's t	-0.901	945	0.368	-0.288	0.320	Cohen's d	-0.0586
	Mann-Whitney U	108605		0.450	-2.92e-5		Rank biserial correlation	0.0284

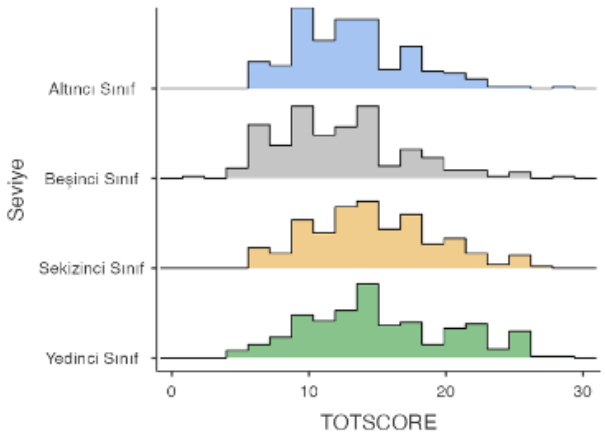
Not. $H_a\ \mu_{KIZ} \neq \mu_{ERKEK}$

Table 15.
Descriptive statistics

									Shapiro-Wilk	
	Grade level	N	Missing	Mean	Median	SS	Min.	Max.	W	p
TOTSCORE	5th grade (Beşinci sınıf)	207	0	12.2	11	4.68	1	28	0.965	<.001
	6th grade (Altıncı sınıf)	222	0	13.0	12.0	4.13	6	28	0.960	<.001
	7th grade (Yedinci sınıf)	288	0	15.2	14.0	5.35	4	28	0.971	<.001
	8th grade (Sekizinci sınıf)	230	0	14.7	14.0	4.62	6	27	0.978	0.001

Figure 8 displays histograms illustrating the distribution of total scores for each grade level.

Figure 8.
Distribution of total scores of the groups by grade levels (histogram)



The results of the analyses for the normality assumption and the homogeneous distribution of variances were presented in Tables 16 and Table 17 below. The results of the Kolmogorov-Smirnov test, which is appropriate for large samples, indicate that the groups are not normally distributed ($p < .001$, $\alpha = .05$).

Table 16.
Normality tests

		Statistics	p
TOTSCORE	Shapiro-Wilk	0.981	<.001
	Kolmogorov-Smirnov	0.0698	<.001
	Anderson-Darling	5.25	<.001

As seen in Table 17, the assumption of homogeneity of variances according to Levene test findings could not be provided, either.

Table 17.
Homogeneity of variances test

		Statistics	df	df2	p
TOTSCORE	Levene	6.40	3	943	<.001
	Bartlett	17.2	3		<.001

Since these assumptions could not be met, the non-parametric Kruskal-Wallis ANOVA test was employed for analysis. The test results presented in Table 18 demonstrate that at least two grade levels differ significantly.

Table 18.
Kruskal-Wallis ANOVA

	χ^2	df	p	ε^2
TOTSCORE	55.9	3	<.001	0.0591

The results of the pairwise comparison tests were displayed in Table 19 in order to determine which two groups are different.

Table 19.
Pairwise comparisons

		W	p
Fifth grade	Sixth grade	-3.02	0.142
Fifth grade	Seventh grade	8.80	<.001
Fifth grade	Eighth grade	7.79	<.001
Sixth grade	Seventh grade	6.66	<.001
Sixth grade	Eighth grade	5.51	<.001
Seventh grade	Eighth grade	1.30	0.793

As seen in Table 19, there is a significant difference between the fifth grades and the seventh and eighth grades in terms of total scores, while there is no significant difference between the fifth and sixth grades in terms of total scores ($p=.142$, $\alpha = .05$). There was no significant difference between the seventh and eighth grades in terms of their total scores ($p=.793$, $\alpha = .05$).

E. Is there a difference between the total scores according to the components of spatial reasoning?

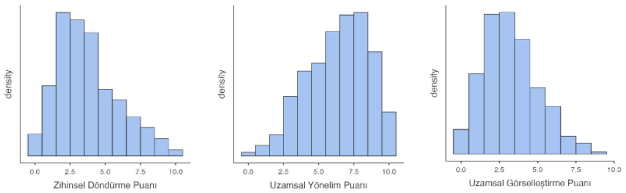
When we analyze the SRT scores according to the sub-components of spatial reasoning, we obtain descriptive statistics for each student's mental rotation score (MRS), spatial orientation score (SOS), and spatial visualization score (SVS) in Table 20 below.

Table 20.
Descriptive statistics

								Shapiro-Wilk	
	N	Missing	Mean	Median	SD	Min.	Max.	W	p
MRS	947	0	3.73	3	2.19	0	10	0.941	<.001
SOS	947	0	6.38	7	2.18	0	10	0.958	<.001
SVS	947	0	3.24	3	1.80	0	9	0.952	<.001

Figure 9 below presents the histograms illustrating the distributions of MRS, SOS and SVS scores.

Figure 9.
Distributions of MRS, SOS and SVS scores (histogram)



The Friedman test was used instead of repeated measures ANOVA for the non-parametric groups to determine whether there is a statistically significant difference between the arithmetic means of the

scores obtained from the three components; the results are presented in Table 21. Separate comparisons of this group's MRS, SOS, and SVS reveal a statistically significant difference (p <.001, α =.05). According to the representations in Figure 9, MRS and SVS have a similar distribution and the mean is lower than SOS.

Table 21.
Friedman

χ^2	df	p
1009	2	<.001

In Table 22, pairwise comparisons of these three scores reveal statistically significant differences between any pair of scores. The values in Table 22 validate the representations in Figure 9.

Tablo 22.
Pairwise comparisons (Durbin-Conover)

			Statistics	p
Mental rotation score (MRS)	-	Spatial orientation score (SOS)	36.21	<.001
Mental rotation score (MRS)	-	Spatial visualization score (SVS)	7.10	<.001
Spatial orientation score (SOS)	-	Spatial visualization score (SVS)	43.32	<.001

F. Is there a gender difference in the total scores derived from the components?

The difference between male and female students' scores on MRS, SOS, and SVS was investigated. First, Table 23 presents the descriptive statistics of the groups.

Table 23.
Group descriptive statistics

	Group	N	Mean	Median	SD	SE
MRS	MALE	499	3.90	3.00	2.20	0.0983
	FEMALE	448	3.54	3.00	2.16	0.1021
SOS	MALE	499	6.68	7.00	2.13	0.0954
	FEMALE	448	6.06	6.00	2.18	0.1032
SVS	MALE	499	3.10	3.00	1.77	0.0792
	FEMALE	448	3.39	3.00	1.82	0.0859

First, the normality test and the Levene test were used to determine whether the groups displayed a normal distribution and whether the variances were homogeneously distributed, respectively. The outcomes of these tests are presented in Tables 24 and 25 below.

Table 24.
Normality tests

		Statistics	p
Mental Rotation Score (MRS)	Shapiro-Wilk	0.954	<.001
	Kolmogorov-Smirnov	0.1231	<.001
	Anderson-Darling	13.86	<.001
Spatial Orientation Score (SOS)	Shapiro-Wilk	0.973	<.001
	Kolmogorov-Smirnov	0.0939	<.001
	Anderson-Darling	7.47	<.001
Spatial Visualization Score (SVS)	Shapiro-Wilk	0.969	<.001
	Kolmogorov-Smirnov	0.1206	<.001
	Anderson-Darling	9.18	<.001

The Kolmogorov-Smirnov test results presented in Table 24 demonstrate that the assumption of normality could not be met.

Table 25.
Homogeneity of variances test

		Statistics	df	df2	p
Mental Rotation Score (MRS)	Levene's	0.0199	1	945	0.888
	Bartlett's	0.124	1		0.724
Spatial Orientation Score (SOS)	Levene's	0.3409	1	945	0.559
	Bartlett's	0.282	1		0.595
Spatial Visualization Score (SVS)	Levene's	0.8563	1	945	0.355
	Bartlett's	0.366	1		0.545

Table 25 demonstrates that the variances are not distributed homogeneously. Accordingly, the Kruskal-Wallis test was conducted on non-parametric data sets, and the outcomes are presented in Table 26 below.

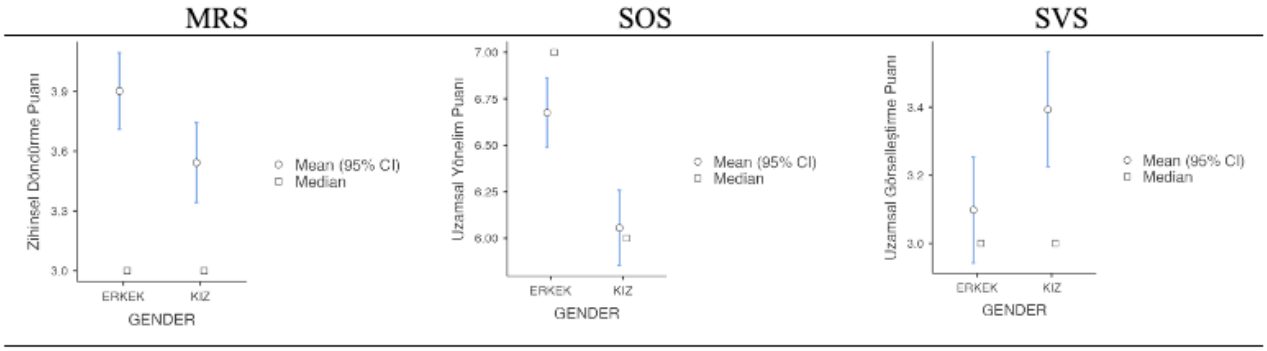
Table 26.
Kruskal-Wallis ANOVA

	χ^2	df	p	ε^2
Mental Rotation Score (MRS)	5.88	1	0.015	0.00622
Spatial Orientation Score (SOS)	19.55	1	<.001	0.02066
Spatial Visualization Score (SVS)	5.39	1	0.020	0.00570

We can conclude that there is a statistically significant difference between female and male students' MRS, SOS, and SVS based on Table 26's p values (respectively p= .015, p < .001, p = .020). The difference between male and female pupils is depicted visually in Figure 10 below.

As shown in Figure 10, there is a statistically significant difference between male and female students MRS, SOS, and SVS scores. Although male students score higher on mental rotation and spatial orientation, female students score higher on spatial visualization.

Figure 10.
Distribution of students' MRS, SOS and SVS by gender



Discussion and Conclusion

SRT, which was developed by Ramful et al. (2017) to measure spatial reasoning ability with its three sub-components, was shown to obtain valid and reliable results when translated into Turkish. The construct validity of the test was shown by the correlation values between the MRS, SOS, SVS, and total scores and the values obtained from confirmatory factor analysis. Furthermore, the test can be claimed to be valid for mental rotation, spatial orientation, and spatial visualization components.

When analyzing the descriptive statistics of the total scores of 947 secondary school students to whom SRT was administered, the mean is low. When compared to the mean values of Ramful et al. (2017), the lower average (13.4) suggests that students' spatial reasoning abilities may be improved. This result indicates that spatial reasoning skills in middle schools should be developed through the use of alternative teaching techniques and strategies or innovative approaches (Baki, 2000; Cheng & Mix, 2014; Lowrie & Logan, 2018; Yolcu, 2008).

Gender variable

According to the study's findings, there was no significant difference between male and female students' total spatial reasoning scores, but there was a significant difference between male and female students' spatial reasoning subcomponent scores. Numerous early research using spatial reasoning abilities tests showed that males outperformed females (Harris, 1978; Maccoby & Jacklin, 1974). However, this study's conclusion that there is no significant gender difference in spatial reasoning abilities is consistent with the findings in several recent studies (Kaya, 2019; Turğut & Yılmaz, 2012; Uzun, 2019). Bartlett & Camba (2023) found that, as the number of studies measuring spatial reasoning has increased, the variation of spatial reasoning ability by gender in the results of the tests used in the majority of studies measuring spatial reasoning has become very small and insignificant. The difference between genders emerges in the sub-components of spatial reasoning skill in the findings reported here. It was determined, for instance, that male students performed better on the mental rotation subcomponent. This finding is consistent with the findings of Voyer and Doyle (2010) and Turğut et al. (2017). When Bartlett and Camba (2023) analyzed the studies that demonstrated a significant difference based on the gender variable and the tests used in these studies, they offered various explanations for this phenomenon. It is claimed that the alleged difference may be due to the fact that the test used was constructed in a way that makes a difference according to the gender variable, the development of gender roles, and spatial reasoning skills (Connell, 2021); the difference found in early studies may be due to the fact that other studies have been cited for years and have influenced these studies (Bartlett & Camba, 2023). In addition, the results of previous studies can be explained by factors such as the society's changing perspective on gender roles and spatial reasoning, which acts as a self-fulfilling prophecy (Bartlett & Camba, 2023). In addition, Moè (2009) found that female students who was convinced they were capable at mental rotation performed better than male students who believed rotation to be difficult. It has been suggested that female students' beliefs about mental rotation have no effect on their performance on the spatial rotation subcomponent, whereas male students' beliefs have a negative effect on their performance (Moè, 2009).

In addition to scoring lower on mental rotation, female students also scored lower on the spatial orientation component. Males have lower spatial visualization scores than females. While there is no statistically significant difference between genders in terms of total score, the difference between subcomponents suggests that these components alone are insufficient to explain spatial reasoning ability (Bartlett & Camba, 2023). In addition, it can be argued that this finding may be the result of cognitive development differences between individuals of different genders, which may also be reflected in the subcomponents of spatial reasoning abilities.

Grade level variable

According to the findings, there was a significant difference in total scores between the fifth and seventh and eighth grades, but there was no significant difference between the fifth and sixth grades and the seventh and eighth grades. In this case, we can argue that the fifth and sixth grades are at a comparable level, as are the seventh and eighth grades. In addition, the total test scores of students in the seventh and eighth grades are substantially higher than those of students in the fifth and sixth grades. Early childhood education and living experiences impact spatial skills (Akkaya Yılmaz, 2022; Turğut, 2007). From this perspective, it is reasonable to anticipate that the student's spatial reasoning skills will improve as the level progresses and will continue to improve over time. Examining the research's findings reveals that spatial reasoning skills improve with increasing grade level. This finding is consistent with the outcomes of prior research (Turğut & Yılmaz, 2012; Akkaya Yılmaz, 2022). The inclusion of learning objectives requiring spatial reasoning skills in the middle school mathematics curriculum, the increase in the number of learning objectives, and the consequent increase in the time allocated to the development of spatial reasoning skills may all contribute to the development of this skill. We can also explain that questions requiring spatial reasoning are encountered more frequently by 7th and 8th grade students during the high school transition exam period, whose spatial reasoning skill scores are anticipated to increase as the level progresses. In addition, the fact that spatial reasoning skill is associated more with other mathematical concepts such as coordinate plane, line, equation, and geometric objects in the 7th and 8th grade mathematics curriculum and the fact that various mathematical concepts are learned as the grade level increases (Akkaya Yılmaz, 2022) play a significant role in this outcome. In addition, Ben-Chaim et al. (1988)'s claim that the seventh-grade level is ideal for developing spatial reasoning may have contributed to this result. It can be said that it is essential to increase the teaching environments and opportunities that contribute to the development of spatial reasoning skills, as well as the activities that support them. From this perspective, it can be argued that it is essential to devote more time to activities aimed at developing spatial reasoning skills in younger age groups and that it is necessary to design learning environments to foster the development of spatial reasoning skills (Turğut & Yılmaz, 2017).

Components of spatial reasoning skill

The results of pairwise comparisons (MRS-SOS, MRS-SVS, SOS-SVS) of the total scores derived from the spatial reasoning skill components indicate that there is a statistically significant difference between each pair. This result is consistent with Ramful et al. (2017)'s findings. The average scores for mental rotation and spatial visualization are lower than those for spatial orientation. Although mental rotation and spatial orientation demand similar skills, mental rotation requires the movement of the object and the mental reconstruction of its elements, so it is similar to spatial orientation in this regard (Turğut & Yılmaz, 2012). The result of this study supports this conclusion. In addition, it was determined that the cumulative scores for each of these components were related to one another and to the overall score.

Implications

The findings of this study can be assessed by math teachers and curriculum designers. Teachers of mathematics can take into consideration the components highlighted here when selecting appropriate activities for acquiring spatial reasoning skills and can determine their students' level by applying SRT. Teachers can take a deliberate approach to selecting and designing the most appropriate teaching environment, knowing that the development of spatial reasoning skills is critical to their students' success (Sorby & Panther, 2020) in national exams such as high school entrance exams or international exams such as PISA and TIMSS.

Specialist faculty members assigned by the Ministry of National Education as one of the groups that design or develop the mathematics curriculum can also precisely target each component of spatial reasoning skills, ascertain the appropriate learning objectives, and propose relevant activities. The fact that spatial reasoning skill is closely related to mathematics achievement (Mix & Cheng, 2012) is further evidence of the need to investigate its origins, nature, and scope. Future research may concentrate on identifying these connections and demonstrating how and at what level they exist.

References

Anggraini, S., Setyaningrum, W., & Retnawati, H. (2020). How to improve critical thinking skills and spatial reasoning with augmented reality in mathematics learning? In *Journal of Physics: Conference Series* (Vol. 1581, No. 1, p. 012066). IOP Publishing.

Akkaya Yılmaz, M. (2022). Spatial reasoning skills levels of junior high school students. *International Journal of Geography and Geography Education (IGGE) (47)*, 135-147. http://dx.doi.org/10.32003/igge.1116462

Atasoy, B., Yüksel, A.O. & Özdemir, S. (2019). Impact of 3D design on spatial ability: Hackidhon case. *Gazi University Journal of Education Faculty, 39*(1), 341-371.

Baki, A. (2000). Learning mathematics within a computer-based environment. *Hacettepe University Journal of Education Faculty, 19*, 186-193.

Baki, A., Kosa, T., & Güven, B. (2011). A comparative study of the effects of using dynamic geometry software and physical manipulatives on the spatial visualisation skills of pre-service mathematics teachers. *British Journal of Educational Technology, 42*(2), 291-310.

Bartlett, K. A., & Camba, J. D. (2023). Gender differences in spatial ability: A critical review. *Educational Psychology Review*, *35*(1), 8.

Battista, M. T. (1990). Spatial visualization and gender differences in high school geometry. *Journal for Research in Mathematics Education, 21*(3), 47-60.

Battista, M. T., Frazee, L. M., & Winer, M. L. (2018). Analyzing the relation between spatial and geometric reasoning for elementary and middle school students. In *Visualizing mathematics* (pp. 195-228). Springer, Cham.

Ben-Chaim, D., Lappan, G., & Houang, R. T. (1988). The effect of instruction on spatial visualization skills of middle school boys and girls. *American Educational Research Journal, 25*(1), 51-71.

Burte, H., Gardony, A. L., Hutton, A., & Taylor, H. A. (2017). Think3d! Improving mathematics learning through embodied spatial training. *Cognitive Research: Principles and Implications, 2*, 1-18.

Callingham, R., Bond, T. (2006). Research in mathematics education and rasch measurement. *Mathematics Education Research Journal, 18*(2), 1–10. https://doi.org/10.1007/BF03217432

Casey, B. M., Andrews, N., Schincler, H., Kersh, J. E., Samper, A., & Copley, J. (2008). The development of spatial skills through interventions involving block building activities. *Cognition and Instruction, 26*(3), 269–309.

Cheng, Y. L., & Mix, K. S. (2014). Spatial training improves children's mathematics ability. *Journal of cognition and development, 15*(1), 2-11.

Christensen, K. B., Makransky, G., & Horton, M. (2017). Critical values for Yen's Q 3: Identification of local dependence in the Rasch model using residual correlations. *Applied psychological measurement, 41*(3), 178-194. doi: 10.1177/0146621616677520.

Clements, D. H. (1998). Geometric and spatial thinking in young children, State University of New York at Buffalo. *Mathematics for the Young Children*, 3-29.

Clements, D. H., & Battista, M. T. (1992). Geometry and spatial reasoning. In D. A. Grouws (Ed.), *Handbook of research on mathematics teaching and learning* (pp. 420–464). New York: Macmillan.

Clements, D. H., & Sarama, J. (2011). Early childhood teacher education: The case of geometry. *Journal of Mathematics Teacher Education, 14*(2), 133–148.

Connell, R. (2021). *Gender in World Perspective (Fourth Edition)*. Polity Press.

de Ayala, R. J. (2009). *The theory and practice of item response theory*. The Guilford Press.

Fowler, S., Cutting, C., Kennedy, J., Leonard, S. N., Gabriel, F., & Jaeschke, W. (2022). Technology enhanced learning environments and the potential for enhancing spatial reasoning: a mixed methods study. *Mathematics Education Research Journal, 34*(4), 887-910.

Fujita, T., Kondo, Y., Kumakura, H., Kunimune, S., & Jones, K. (2020). Spatial reasoning skills about 2D representations of 3D geometrical shapes in grades 4 to 9. *Mathematics Education Research Journal, 32*, 235-255.

Güven, B., & Kösa, T. (2008). The effect of dynamic geometry software on student mathematics teachers' spatial visualization skills. *Turkish Online Journal of Educational Technology-TOJET, 7*(4), 100-107.

Gün, E., & Atasoy, B. (2017). The effects of augmented reality on elementary school students' spatial ability and academic achievement. *Education and Science 42*(191), 31-51.

Guillot, A., Hoyek, N., Louis, M., & Collet, C. (2012). Understanding the timing of motor imagery: Recent findings and future directions. *International Review of Sport and Exercise Psychology, 5*(1), 3–22. https ://doi.org/10.1080/17509 84X.2011.62378 7.

Gravetter, F. J., & Forzano, L. A. B. (2018). *Research methods for the behavioral sciences*. Cengage Learning.

Harris, L. J. (1978). Sex differences in spatial ability: Possible environmental, genetic, and neurological factors. In M. Kinsbourne (Ed.), *Asymmetrical Function of the Brain.* Cambridge University Press.

İrioğlu, Z. ve Ertekin, E. (2012). Investigation of mental rotation skills of middle school students in terms of some variables [İlköğretim ikinci kademe öğrencilerinin zihinsel döndürme becerilerinin bazı değişkenler açısından incelenmesi]. *Journal of Educational and Instructional Studies in the World, 2*(1).

Kabakçı, D. A. ve Demirkapı, A. (2016). The effect of "mathematics and art" course activity applications on students' spatial talents in Izmit science and art center. *HAYEF Journal of Education, 13*(1), 11-22.

Kaya, D. (2019). Relationship between spatial visualization, mental rotation and mental visualization skills with mathematics-oriented epistemological beliefs and some other variables of eighth grade students. *Kastamonu Education Journal*, *27*(4), 1787-1798.

Kayhan, E. B. (2005). *Investigation of high school student' spatial ability.* [Unpublished doctoral dissertation]. Middle East Technical University.

Linacre, J. M. (2016). *Winsteps® Rasch measurement computer program User's Guide*. Beaverton, Oregon: Winsteps.com

Lohman, D. F. (1979). *Spatial ability: A review and reanalysis of the correlational literature* (Tech. Rep. No. 8), Stanford, CA: Stanford University, Aptitude Research project, School of Education. (NTIS NO. AD-A075 972).

Lohman, D. F. (1996). Spatial ability and G. In I. Dennis & P. Tapsfield (Eds.), *Human abilities: Their nature and assessment* (pp. 97–116). Hillsdale, NJ: Lawrence Erlbaum.

Lowrie, T., & Logan, T. (2018). The interaction between spatial reasoning constructs and mathematics understandings in elementary classrooms. In *Visualizing mathematics* (pp. 253-276). Springer, Cham.

Lowrie, T., Logan, T., & Hegarty, M. (2019). The influence of spatial visualization training on students' spatial reasoning and mathematics performance. *Journal of Cognition and Development, 20*(5), 729–751.

Maccoby, E. E., & Jacklin, C. N. (1974). *The Psychology of Sex Differences*. Stanford University Press.

McGee, M. G. (1979) Human spatial abilities: Psychometric studies and environmental, genetic, hormonal, and neurological influences. *Psychological Bulletin, 86*(5), 889-911.

Ministry of National Education (MoNE) [Milli Eğitim Bakanlığı (MEB)], (2018). Middle School Mathematics (5th, 6th, 7th and 8th Grades) Curriculum. Ankara: Milli Eğitim Basımevi.

Memişoğlu, B., & Tapan-Broutin, M. S. (2018). Evaluation of the conceptual geometry concepts from the republican today mathematics education programs. *Uluslararası Bilimsel Araştırmalar Dergisi (IBAD)*, 196-209.

Mix, K. S., & Cheng, U. L. (2012). The relation between space and math: Developmental and educational implications. In J. B. Benson (Ed.), *Advances in child development and behavior* (Vol. 42, pp. 199–243). Waltham, MA: Academic Press.

Moè, A. (2009). Are males always better than females in mental rotation? Exploring a gender belief explanation. *Learning and Individual differences*, *19*(1), 21-27.

Mulligan, J. (2015). Looking within and beyond the geometry curriculum: connecting spatial reasoning to mathematics learning. *ZDM, 47*, 511-517.

NCTM (2000). Principles and standards for school mathematics. Reston, VA: National Council of Teachers of Mathematics.

Okagaki, L., & Frensch, P. A. (1994). Effects of video game playing on measures of spatial performance: Gender effects in late adolescence. *Journal of applied developmental psychology, 15*(1), 33-58.

Olkun, S. (2003). Making connections: Improving spatial abilities with engineering drawing activities. *International journal of mathematics teaching and learning, 3*(1), 1-10.

Pietropaolo, S., & Crusio, W. E. (2012). Learning spatial orientation. In N. M. Seel (Ed.), *Encyclopedia of the Sciences of Learning* (pp. 1969–1971). Boston, MA: Springer, US.

Ramful, A., Lowrie, T., & Logan, T. (2017). Measurement of spatial ability: Construction and validation of the spatial reasoning instrument for middle school students. *Journal of Psychoeducational Assessment, 35*(7), 709-727.

Resnick, I., Newcombe, N. S., & Jordan, N. C. (2019). The relation between spatial reasoning and mathematics achievement in children with mathematics learning difficulties. In P. Rasanen (Ed.), International handbook of mathematical learning difficulties. Springer.

Resnick, I., Harris, D., Logan, T., & Lowrie, T. (2020). The relation between mathematics achievement and spatial reasoning. *Mathematics Education Research Journal, 32*, 171-174.

R Core Team (2022). R: A Language and environment for statistical computing. (Version 4.1) [Computer software]. Retrieved from https://cran.r-project.org. (R packages retrieved from CRAN snapshot 2023-04-07).

Seng, S., & Chan, B. (2000). Spatial ability and mathematical performance: Gender difference in an elementary school. (Report No. PS028362). Singapore: National Institute of Education, Nanyang Technological University. (ERIC Document Reproduction Service No. ED438937)

Seol, H. (2023). snowRMM: Rasch Mixture, LCA, and Test Equating Analysis. (Version 5.5.0) [jamovi module]. URL https://github.com/hyunsooseol/snowRMM.

Shea, D. L., Lubinski, D., and Benbow, C. P. (2001). Importance of assessing spatial ability in intellectually talented young adolescents: a 20-year longitudinal study. *Journal of Educational Psychology, 93*, 604–614. doi: 10.1037/0022-0663.93.3.604

Sinclair, N., & Bruce, C. (2015). New opportunities in geometry education at the primary school. *ZDM – The International Journal on Mathematics Education, 47*(3), 319.

Sorby S. (1999). Developing 3-D spatial visualization skills. *Engineering Desing Graphics Journal, 63*(2), 21-32.

Sorby, S. A., & Panther, G. C. (2020). Is the key to better PISA math scores improving spatial skills? *Mathematics Education Research Journal, 32*(2), 213-233.

Tabachnick, B. G., & Fidell, L. S. (2019). Using multivariate statistics (Seventh Edition). Pearson.

The Jamovi project (2023). *Jamovi*. (Version 2.4) [Computer Software]. Retrieved from https://www.jamovi.org.

Turğut, M. (2007). *Investigation of 6, 7 and 8 grade students' spatial ability.* [Unpublished master's thesis]. Dokuz Eylül University.

Turğut, M. & Yılmaz, S. (2012). Investigation of 7th and 8th grade students' spatial ability *Dicle University Ziya Gökalp Journal of Education Faculty, 21*, 69-79.

Turğut, M., Yenilmez, K. & Balbağ, M. Z. (2017). Prospective teachers' logical and spatial thinking skills: The effects of department, gender and academic performance. *Mehmet Akif Ersoy University Journal of Education Faculty 1*(41), 265-283.

Turhan, E. İ. (2010). *The effect of computer aided perspective drawings on eighth grade students' spatial abilities and attitudes towards mathematics, technology and geometry.* [Unpublished doctoral dissertation]. Eskisehir Osmangazi University, Eskişehir.

Uttal, D. H., Meadow, N. G., Tipton, E., Hand, L. L., Alden, A. R., Warren, C., & Newcombe, N. S. (2013). The malleability of spatial skills: a meta-analysis of training studies. *Psychological Bulletin, 139*(2), 352–402.

Uygan, C. (2011). *The effect of google sketchup and concrete model supported applications on the spatial abilities of elementary mathematics teacher candidates in teaching solid objects.* [Unpublished master's thesis]. Eskişehir Osmangazi University.

Uygan, C. & Turğut, M. (2012). Examination of mathematics questions involving the use of spatial ability in national central examinations [Ulusal merkezi sınavlarda uzamsal yeteneğin kullanımını içeren matematik sorularının incelenmesi]. X. Fen bilimleri ve Matematik eğitimi kongresi bildiri özetleri kitabı [10th Science and Mathematics Education Congress Proceedings Book], 10, 33.

Uzun, Z. B. (2019). *Middle school students geometric thinking levels, spatial abilities and attitudes towards geometry.* [Unpublished master's thesis]. Balıkesir University.

Van de Walle, J. A, Karp, K. S. & Bay Williams, J. M. (2012). İlkokul ve ortaokul matematiği gelişimsel yaklaşımla öğretim. (Editör: Soner Durmuş, Çevirmen: Yüksel Dede. Ankara: Nobel Yayın Dağıtım.

Voyer, D., & Doyle, R. A. (2010). Item type and gender differences on the Mental Rotations Test. *Learning and Individual Differences, 20*(5), 469-472.

Wai, J., Lubinski, D., & Benbow, C. P. (2009). Spatial ability for STEM domains: Aligning 50 years of cumulative psychological knowledge solidifies its importance. *Journal of Educational Psychology, 101*, 817-835.

Woolcott, G., Le Tran, T., Mulligan, J., Davis, B., & Mitchelmore, M. (2022). Towards a framework for spatial reasoning and primary mathematics learning: an analytical synthesis of intervention studies. *Mathematics Education Research Journal, 34*, 37-67.

Yen, W. M. (1984). Effects of local item dependence on the fit and equating performance of the three-parameter logistic model. *Applied Psychological Measurement, 8*(2), 125-145.

Yıldırım Gül, Ç. & Karataş, İ. (2015). Investigation of correlation among the 8th grade students' achievement on transformation geometry, spatial ability, levels of geometry understanding and attitudes towards mathematics. *Karaelmas Journal of Educational Sciences, 3(1)*, 36-48.

Yıldız B. (2009). *The effects of using three-dimensional virtual environments and concrete manipulatives on spatial visualization and mental rotation abilities*. [Unpublished master's thesis]. Hacettepe University.

Yıldız, B., & Tüzün, H. (2011). Effects of using three-dimensional virtual environments and concrete manipulatives on spatial ability. *Hacettepe University Journal of Education, 41*(41).

Yolcu, B. (2008). *The study of improving the spatial ability of sixth grade students with concrete models and computer practicing*. [Unpublished master's thesis]. Eskişehir Osmangazi University.

Yurt, E., & Sünbül, A. M. (2012). Effect of modeling-based activities developed using virtual environments and concrete objects on spatial thinking and mental rotation skills. *Educational Sciences: Theory & Practice, 12*(3), 1975-1992.