

Article

Not peer-reviewed version

Artificial Intelligence Applications in Primary Education: A Quantitatively-Supported Mixed-Meta Method Study

Yavuz Topkaya , [Yunus Doğan](#) ^{*} , Veli Batdı , [Sami Aydın](#)

Posted Date: 30 January 2025

doi: [10.20944/preprints202501.2263.v1](https://doi.org/10.20944/preprints202501.2263.v1)

Keywords: Artificial Intelligence; Artificial Intelligence in Education; Mixed-methods



Preprints.org is a free multidisciplinary platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This open access article is published under a Creative Commons CC BY 4.0 license, which permit the free download, distribution, and reuse, provided that the author and preprint are cited in any reuse.

Article

Artificial Intelligence Applications in Primary Education: A Quantitatively-Supported Mixed-Meta Method Study

Yavuz Topkaya ¹, Yunus Doğan ^{2,*}, Veli Batdı ³ and Sami Aydın ⁴

¹ Mustafa Kemal University, Hatay, Türkiye

² Fırat University, School of Foreign Languages, Elazığ-Türkiye

³ Gaziantep University, Department of Curriculum and Instruction, Gaziantep-Türkiye

⁴ Gaziantep University Education Faculty, Gaziantep/Türkiye

* Correspondence: jonah.saidson@gmail.com

Abstract: Education systems are undergoing radical transformations under the influence of digital transformation. One of the leading actors in this digital transformation is Artificial Intelligence (AI) applications. This study evaluates the effects of various variables related to AI applications in primary education using a mixed-meta method integrated with quantitative analysis. As part of the mixed-meta method, a meta-analysis of data from studies conducted between 2005 and 2025 was performed using the CMA software. The analysis revealed a moderate effect size of $g = 0.51$. To ensure the validity of the meta-analysis results and enhance their scope, a meta-thematic analysis was conducted, where content analysis identified themes and codes. In the final phase of the study, an evaluation form prepared within the framework of the Rasch measurement model was administered to primary school teachers to support the findings obtained from the mixed-meta method. The data collected were analyzed using the FACETS software. The study's results indicated that, in the meta-analysis of documents, AI studies were predominantly conducted in mathematics education at the primary level. In the meta-thematic analysis, themes emerged regarding the impact of AI applications on learning environments, challenges encountered during implementation, and proposed solutions. In the Rasch measurement model process, it was observed that AI applications were widely used in science and mathematics curricula (FBP-4 and MP-2). Among the evaluators (raters), J2 was the most lenient scorer, while J11 was the strictest. Regarding the items on the use of AI applications, the statement "I can enable students to prepare a presentation describing their environment using AI tools" (I17) was identified as the most challenging, whereas the statement "I know how to use AI applications effectively in classroom activities" (I14) was the easiest. The findings indicate that the data obtained from the different analysis processes complement and support each other. It is anticipated that the study's results will guide future research and applications in this field and make significant contributions to the domain.

Keywords: quantitative-supported mixed-meta method; meta-analysis; meta-thematic analysis; rasch measurement model; Maxqda

1. Introduction

By the end of the 20th century, advancements in technology and technical fields caused profound changes in societies. Today, information technologies lie at the heart of constantly evolving education systems (Candan & Başaran, 2023). During this period, one of the most notable developments among existing technologies is artificial intelligence. Artificial intelligence can be defined as a technology capable of mimicking certain abilities of computer systems (Russell & Norvig, 2009). This technology, characterized as machines that can autonomously think, analyze large datasets, recognize patterns, learn, and solve problems, has been shaping the future of computer

systems and has become an indispensable part of daily life (Akerkar, 2014; Ginsenberg, 2012). Thanks to the advancing intelligence of computers, continuous innovations in software and hardware are emerging, integrating various technologies like robots, smart home systems, and autonomous vehicles into our daily lives (Komalavalli et al., 2020).

Artificial intelligence has a wide range of applications across fields such as healthcare, finance, manufacturing, and education. For instance, in the healthcare sector, artificial intelligence can analyze hospital data to create diagnostic and treatment plans and provide preventive healthcare services (Topol, 2019). In financial services, artificial intelligence can be used to prevent fraud and enhance customer service (Litan, 2021). In the industrial sector, it offers advantages such as lower-cost and widespread production, along with safer transportation options (Hamet & Tremblay, 2017; Li et al., 2017). The rapid progress of artificial intelligence is deeply transforming our lifestyles, learning methods, and ways of conducting business. The proliferation of innovative artificial intelligence applications has significant implications for society and future generations. Artificial intelligence has moved beyond professional and academic research to become a vital component of everyday life. Thus, artificial intelligence education needs to evolve from being limited to experts and be extended to the general public (Chiu & Chai, 2020).

While artificial intelligence topics are generally addressed in higher education, they have begun to be implemented in primary and secondary school classrooms worldwide as part of a strategic initiative (Pedró et al., 2019). This initiative aims to educate future generations. Throughout artificial intelligence education, extending even to secondary education levels, children not only become aware of emerging technologies and how they work but also receive inspiration to become the future users, designers, software developers, and researchers of artificial intelligence (Pedró et al., 2019).

1.1. Artificial Intelligence in Education

As in many fields, artificial intelligence (AI) is being utilized in education. The significance of applying AI in the education sector became evident during the COVID-19 pandemic, which resulted in a shift from traditional classroom teaching to digital learning methods (Maqbool et al., 2021; Mijwil et al., 2022; Pantelimon et al., 2021). In fact, studies on the use of AI in educational settings began as early as the 1980s (Self, 2016). Over the past five decades, research in the literature has focused on AI-supported educational applications (Bracaccio et al., 2019), language teaching processes in educational environments, systems for evaluating and supporting individual performance (Santos, 2016), and intelligent teaching systems and environments (Roscoe, Walker, & Patchan, 2018).

These studies indicate that using AI in education can provide more effective learning experiences by personalizing learning, helping students discover their abilities, fostering creativity, and reducing teachers' workload (Liang & Chen, 2018; Xue & Li, 2018). Additionally, the literature highlights that AI can support students with special needs, offer content tailored to different learning styles, analyze students' learning processes to focus on their individual needs, and guide them in career planning (Mu, 2019). These findings align with the three paradigms defined by Ouyang and Jiao (2021) for the use of AI in education: guiding, supporting, and integrating students into enhanced practices.

While such findings showcase the benefits of AI for students, studies examining teachers' perceptions of AI reveal that teachers often see AI as a professional threat that could replace their jobs (Luckin et al., 2016). However, recent research indicates that these perceptions are changing, as teachers now expect AI to bring significant advancements to education in various settings (Panigrahi, 2020). Teachers' perceptions of AI systems depend on factors such as their pedagogical beliefs, teaching experience, and history of using educational technology. These factors can influence their willingness to adopt new educational technologies (Ryu & Han, 2018).

Studies exploring teachers' perceptions of AI in education (AIE) reveal that teachers see AI as a tool that can contribute to the teaching-learning process by providing digitized learning materials and facilitating human-computer interaction (Jia et al., 2020). They also perceive AI as a means to address learning difficulties faced by students in overcrowded classrooms (Holmes et al., 2019).

Moreover, research indicates that teachers believe AIE can reduce their administrative workload by taking over repetitive and simple tasks (Qin et al., 2020).

1.2. Artificial Intelligence in Elementary Schools

Artificial intelligence (AI) is becoming increasingly significant in our daily lives, with applications in disaster prevention, agriculture, civil engineering, and healthcare, among others (Avci et al., 2021; Jain et al., 2020; Nichols et al., 2019; Sharma et al., 2020). Educating all age groups, from elementary school students to the elderly, about AI and how to use it is crucial for fostering a society where AI is effectively utilized. AI education should not be limited to higher education but should be integrated into all levels, starting from elementary school through secondary education (Su & Yang, 2022).

In 2019, the United Nations Educational, Scientific and Cultural Organization (UNESCO) evaluated educational programs provided by governments and institutions, adopting a policy to support research aimed at developing standardized AI curricula for K-12 education (UNESCO, 2019). This initiative has led to an increase in research on K-12 education and the emergence of conferences and journals dedicated to the topic (Yue et al., 2022).

As children increasingly interact with new technological tools and engage with learning processes through AI systems, the importance of AI education tailored to children is growing. For instance, Francis et al. (2021) found significant improvements in elementary students' spatial reasoning skills after participating in short- or long-term robotic interventions.

Chu et al. (2014) developed a Model-Tracing Intelligent Tutor (MIT) to interpret students' mathematical problem-solving behaviors, diagnose learning difficulties, and provide individual feedback. Their findings showed that the model-tracing-based learning diagnosis system significantly outperformed traditional web-based tests in helping students learn mathematics.

Calero et al. (2019) investigated the impact of robotics-based education on students' mental rotation abilities. Their study revealed that robotics-based education significantly improved boys' mental rotation skills compared to the control group, while no significant difference was observed for girls. Although the study provides evidence of the potential of robotics-based education, it also highlighted the need for future research to deeply analyze gender differences in learning outcomes achieved through educational robotics.

Yang (2022) emphasized the importance of introducing children to AI concepts at an early age and proposed a curriculum design focusing on why, what, and how to teach AI. Similarly, Holmes et al. (2019) echoed this perspective, underscoring the value of early AI education for children.

1.3. Purpose and Significance of the Study

The 21st century has witnessed revolutionary transformations in technology, reshaping the landscape of education. These changes have shifted the acquisition of knowledge beyond individual memory or traditional classroom settings, incorporating interactions through online networks and communities. Consequently, the learning process is now conceptualized as a network structure where knowledge is constantly updated by connecting various nodes of information.

The connectivism theory emerges within this context, emphasizing its effectiveness in fields where knowledge evolves rapidly and requires continuous updating, such as technology and digital literacy. Connectivism perceives learning as a multi-layered process occurring across networks (Siemens, 2004). In other words, it aims to redefine teaching and learning processes in knowledge societies shaped by emerging technologies. Furthermore, connectivism is particularly relevant in explaining learning within e-learning environments (Goldie, 2016) and contributes to the modernization of traditional education by integrating e-learning scenarios (Guerra, 2022). In light of these perspectives, connectivism provides significant contributions to educators aiming to support modern educational settings with digital technology, particularly through AI applications. Artificial intelligence (AI) is defined as the capacity of a digital device to perform tasks typically attributed to humans (Chiu, 2021). In today's era, innovative technologies like AI are increasingly playing a critical

role in classroom environments and students' learning processes (Humble & Mozelius, 2019; Kaplan-Rakowski et al., 2023). AI-related studies focus on various aspects, including enhancing student learning, supporting teaching practices, evaluating student performance, and improving school administration (Chiu et al., 2023). Among these, research on student learning has centered on academic performance, 21st-century skills, and student motivation and engagement (Fu et al., 2020; Li & Su, 2020; Luo, 2018).

When viewed holistically, it becomes evident that there is a gap in studies examining the effects of AI applications on various variables in elementary education, particularly using a mixed-meta approach that integrates quantitative analysis. This study aims to explore the influence of AI applications on different variables in three dimensions;

1. **Quantitative dimension:** Meta-analysis
2. **Qualitative dimension:** Meta-thematic analysis
3. **Quantitative teacher feedback:** Rasch measurement model

The study intends to provide a unique perspective to the literature by integrating these methods. Accordingly, in line with a document analysis framework based on a mixed-meta methodology:

Meta-Analysis

1. Identifying the general effect size of different variables on AI applications.
2. Determining the effect size levels of AI use in terms of courses, implementation duration, and sample size.

Meta-Thematic Analysis

3. Exploring the effects of AI applications on learning environments, identifying potential challenges in the application process, and offering solutions.

Rasch Measurement Model (Teacher Feedback)

4. Analyzing teachers' general opinions on AI applications.
5. Examining jury tendencies towards strictness or leniency in evaluations.
6. Conducting item difficulty analysis for criteria related to AI applications.

2. Method

The methodology of a scientific study encompasses every stage of the research process, from its justification and content to the concepts used, measurement techniques applied, the analysis of collected data, and the interpretation of results (Koçel, 2017). Methodology can be classified into qualitative, quantitative, and mixed-methods approaches (Toraman, 2021). Among these, methodological pluralism, which combines quantitative and qualitative methods, has gained significant traction in fields such as education and sociology (Molina-Azorín et al., 2012).

Methodological pluralism involves the holistic analysis of quantitative and qualitative data through document analysis (Creswell & Sözbilir, 2017). Mixed-methods research, as defined by Creswell et al. (2003), aligns with this concept of methodological pluralism. It involves the integration of qualitative and quantitative approaches, either concurrently or sequentially, to create a unified dataset.

This study adopts a methodological pluralism framework to evaluate the use of AI in primary schools, encompassing both qualitative and quantitative analyses. It employs a mixed-meta approach, which integrates quantitative and qualitative methods within a single research process. Specifically, the methodological process includes:

1. **Meta-analysis:** A quantitative synthesis of data to determine the effect size of AI applications.
2. **Meta-thematic analysis:** A qualitative examination of recurring themes in the literature, focusing on the effects of AI applications in educational contexts.

3. **Rasch measurement model:** A quantitative analysis of participant opinions, providing insights into teacher perspectives and evaluating response consistency.
4. This combination of methods ensures a scientifically robust and holistic research process. The research integrates findings from meta-analysis, meta-thematic analysis, and the Rasch model to comprehensively analyze AI applications in education. A visual representation of this methodological framework is presented in Figure 1.

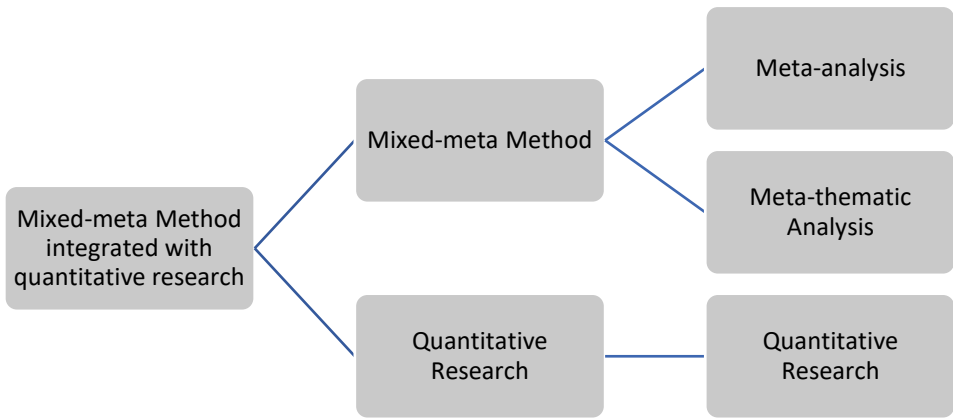


Figure 1. Mixed-meta Method integrated with quantitative research design (Batdı, 2024a).

The mixed-meta method involves the holistic analysis of quantitative and qualitative data based on document examination. In other words, mixed-meta allows for the analysis of quantitative data using statistical programs such as CMA/MetaWin, and qualitative data using software like Nvivo/Maxqda, thereby enabling the integration of both data sets into a single study framework. This makes the method a comprehensive and rich approach to research (Batdı, 2021). The Multiple Complementary Approach (McA), a design within this method, integrates meta-thematic analysis with the quantitative research process to address the shortcomings of meta-thematic findings, support the obtained results, and provide a more holistic perspective (Batdı, 2024b). The analysis of scientifically valuable quantitative and qualitative data is crucial in this research process. In this context, the mixed-meta method integrates meta-analysis, meta-thematic analysis, and Rasch analysis based on participant views in the quantitative dimension. This three-phase approach enhances the depth of the research and ensures the generation of more comprehensive and valid results.

Method Phases:

Meta-Analysis: Statistical analysis of quantitative data, determining the impact of different variables on AI applications.

Meta-Thematic Analysis: Thematic analysis of qualitative data, exploring the impact of artificial intelligence applications on education.

2.1. Meta-Analysis Process

In the first dimension of the research, meta-analysis has been used. Glass (1976) defined the concept of meta-analysis as a statistical analysis process aimed at integrating research findings by combining the analysis results obtained from individual studies. Meta-analysis is a statistical method that, instead of examining individual studies one by one, aims to gather various studies effectively and reliably to produce broader and more meaningful results (Tsagris and Fragkos, 2018). In this dimension of the research, the meta-analysis process combines the general effect size of studies involving artificial intelligence applications at the elementary school level, as well as the effect sizes related to lessons, application duration, and sample size. By doing so, the results of various studies conducted at the elementary school level are effectively and validly integrated, leading to a holistic conclusion (Tsagris and Fragkos, 2018).

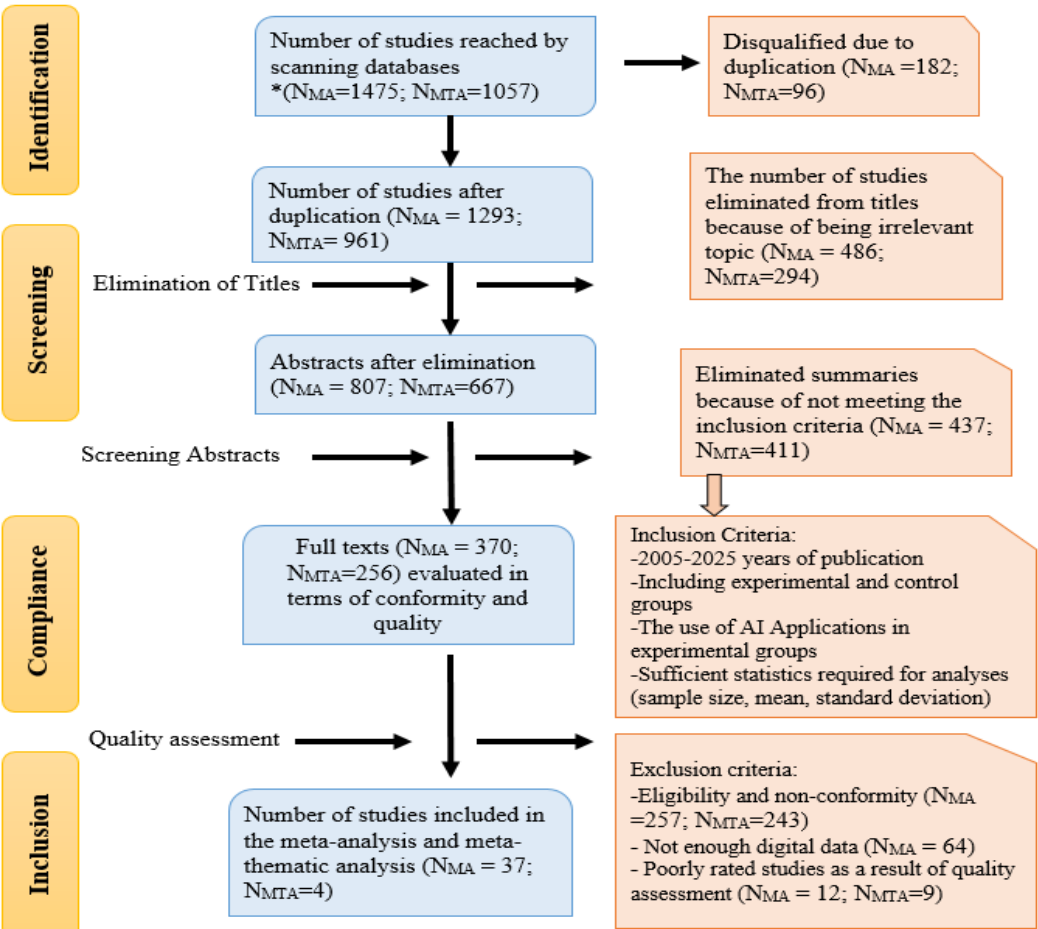
2.1.1. Data Collection and Analysis

Research on the use of artificial intelligence in elementary schools has been conducted in both English and Turkish using the keywords "impact/effectiveness of artificial intelligence use/in elementary schools" in the literature. During this study, databases such as YÖK, Google Scholar, Web of Science, Taylor & Francis Online, Science Direct, and ProQuest Dissertations & Theses Global were searched. The search was conducted according to the inclusion criteria listed below. These criteria are shown in Table 1.

Table 1. Inclusion Criteria for Meta-Analysis Process.

Criteria	Description
Time Period	2005-2025
Publication Language	English and Turkish
Appropriateness of Teaching Method	Experimental and/or quasi-experimental designed studies with pre-test-post-test control groups using artificial intelligence applications
Statistical Data	Sample size (n), arithmetic mean (X), and standard deviation (ss) for effect size calculation

Table 1 outlines the inclusion criteria used to select studies for the meta-analysis. Studies that did not meet these criteria were excluded from the analysis. In this context, studies that lacked access permission, did not contain quantitative data, lacked the necessary data for analysis, were found in multiple databases, or did not involve an experimental process were excluded from the analysis. The number of studies included and excluded, along with the reasons for exclusion, are shown in the PRISMA flow diagram (Moher et al., 2009) in Figure 2.



*Number of the studies for meta-analysis (NMA) and for meta-thematic analysis (NMTA)

Figure 2. Selection of the Studies Included in the Analysis.

As shown in Figure 2, as a result of the screenings, in the first phase, N=1475 studies examining the impact of various variables on the use of artificial intelligence were identified. Out of these studies, 182 were excluded due to duplication, 486 due to irrelevant topics, 437 due to failure to meet inclusion criteria identified through abstract reading (NMA=437), and 64 due to insufficient digital data (NMA=64). As a result, as presented in Figure 2, 12 studies were included in the meta-analysis. Additionally, inter-rater reliability was calculated using the formula $[\text{agreement} / (\text{agreement} + \text{disagreement}) \times 100]$ proposed by Miles and Huberman (1994), and the reliability level of the research was determined to be .90. The data were analyzed using the CMA 2.0 program.

2.1.2. Effect Size and Model Selection

Meta-analysis procedures were carried out using CMA 2.0 software. The effect size (g) obtained from the analyses was interpreted according to the effect levels defined by Thalheimer and Cook (2002). The data were assessed and interpreted within the framework of the random effects model (REM). Schmidt et al. (2009) pointed out that the use of the fixed effects model (SEM) is limited in most cases, emphasizing that REM is a more appropriate option. Therefore, in this research, REM was preferred.

2.1.3. Moderator Analysis

In the meta-analysis of studies on artificial intelligence applications, a heterogeneity test was performed, and an I^2 value of 85.90 was found. This value indicates that the overall effect size of artificial intelligence applications, as well as different variables, could be evaluated. Values with heterogeneity rates of 75% or higher are considered high heterogeneity, so it is recommended to perform moderator analysis (Cooper et al., 2009). Therefore, in this research, to deepen the meta-analysis, various factors influencing artificial intelligence applications were examined, and conducting moderator analysis was deemed necessary.

2.1.4. Publication Bias

In meta-analyses, reliability is a key aspect. In studies where effect size is calculated, the inclusion of published studies or those with significant differences can raise concerns about publication bias. For this reason, certain calculations are made to assess publication bias in meta-analyses. Figure 3 shows the reliability calculations related to this in a funnel plot.

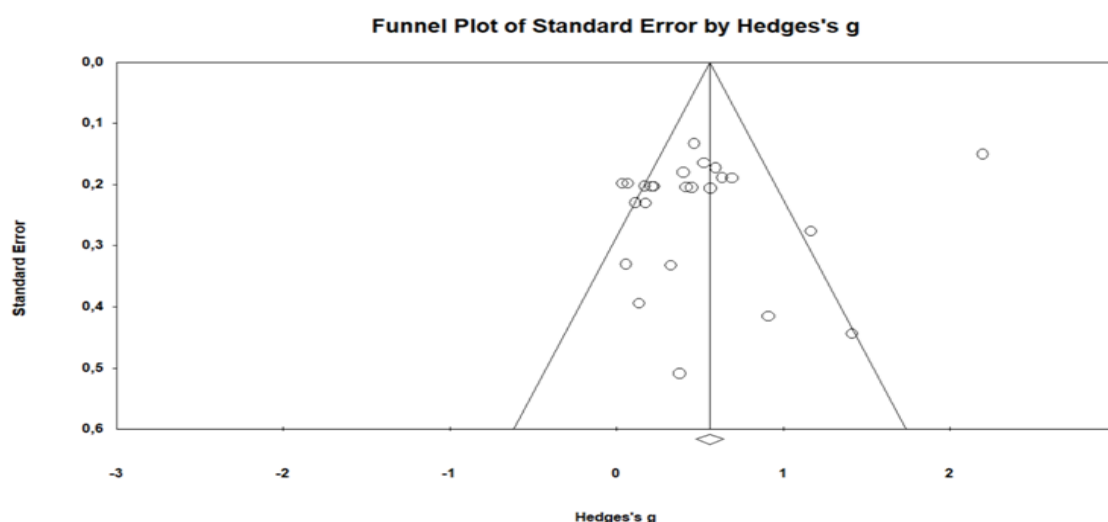


Figure 3. Funnel Plot.

Figure 3 shows a funnel plot, which is presented as a visual summary of the meta-analysis dataset and also illustrates the potential for publication bias (Cooper et al., 2009). Funnel plots are used to detect publication bias in meta-analyses (Duval & Tweedie, 2000). In this plot, the horizontal axis represents effect size, and the vertical axis represents sample size, variance, or standard error values. The funnel plot highlights the relationship between effect size and sample size. As sample size increases, studies tend to cluster closer to the average effect size at the top of the plot (Sterne & Harbord, 2004). If there is no publication bias, the plot should form a symmetrical inverted funnel shape, as seen in Figure 3 (Rodriguez, 2001). Upon examining Figure 3, it can be observed that there is a balanced distribution of studies on both the left and right sides of the symmetry axis. This indicates that no publication bias was found in the research.

To test for publication bias, the calculation of Rosenthal's fail-safe N is a method used for this purpose (Rosenthal, 1979). The N value indicates how many unpublished null studies would be required to invalidate the existing effect. A high N value suggests that the results are valid (Borenstein et al., 2009). In this research, the safe N value is 841. Comparing this with the number of studies included in the analysis, it can be said that this number is quite high, meaning no bias was detected.

2.2. Meta-Thematic Analysis Process

As part of the mixed-meta method, the second dimension of this study applied the meta-thematic analysis process. Meta-thematic analysis can be defined as a type of analysis where participants' views (raw data) from qualitative research on a particular topic are reevaluated, and themes and codes are extracted (Batdı, 2021). In this study, qualitative research on the use of artificial intelligence at the primary school level was reviewed, and the meta-thematic analysis process was evaluated. In meta-thematic analysis, it is not the quantity of data that is important, but rather working with enough data to reach saturation. This type of analysis involves developing new codes and themes from data obtained through document analysis. In other words, meta-thematic analysis is a process based on text or verbal data that combines qualitative findings to generate new themes and codes. The themes and codes obtained in this process are included in the study to provide broader and more comprehensive results. In the qualitative dimension of this study, themes and codes related to the use of artificial intelligence at the primary school level were created using the meta-thematic analysis process (Batdı, 2019).

2.2.1. Data Collection and Review

Qualitative studies on the use of artificial intelligence at the primary school level were gathered using the document review method. These studies were reviewed after a preliminary reading, and the studies that were considered appropriate were selected through four stages: initial reading, full reading, and review of the findings section. The review phase involved reading the title, short summary, and expanded summary before moving to the findings section. After examining the findings section, it was decided whether the study would be included in the current research (Batdı, 2024a). As indicated in the PRISMA diagram, the studies were selected after a four-stage process. Ultimately, four studies were reviewed in the meta-thematic analysis process. Document analysis is defined as the process of collecting, examining, evaluating, and analyzing different documents as primary sources of research data (Sak et al., 2021). In other words, document analysis includes a series of processes involving the examination and evaluation of printed and digital (computer-based and internet-accessible) data (Bowen, 2008). In the processing of meta-thematic data obtained through qualitative research using document analysis, the Maxqda program was used. The data was transferred into the program, and content analysis, a commonly used method (Bowen, 2008), was applied to analyze indicators, comments, and discourses (Bryman, 2016). In this context, the participant opinions obtained in this research were reanalyzed to generate different codes and themes.

2.2.2. Coding Process

An important phase of the meta-thematic analysis process is coding. The opinions of participants were reformulated, and codes were created, grouping similar codes under common themes. In this phase, the findings of the meta-thematic analysis process, which were based on the opinions of participants from the reviewed studies, included the themes and codes determined for the study's main objectives. The reliability of the coding process is demonstrated by ensuring that the generated codes are consistent with each other (Mayring, 2000). In this study, two themes were created at the end of the meta-thematic analysis process: "the impact of AI applications on learning environments" and "problems and solutions encountered in AI applications." Two coders independently created the themes and codes. One of the coders was the researcher, and the other was an academic expert in the field. After completing the coding process, the themes and codes created by both coders were compared to check for consistency and agreement. Similar themes and codes were recorded jointly, while discrepancies were discussed until agreement was reached between the coders. The reliability between coders was calculated using Cohen's Kappa coefficient (Cohen, 1960). The Kappa value for the theme "Impact of AI applications on learning environments" was 0.79, and for the theme "Problems and solutions encountered in AI applications" it was 0.86 (Appendix A). The Kappa values ranging from 0.79 to 0.86 indicate a good to very good level of agreement between the coders.

2.2.3. Reliability in the Meta-Thematic Analysis Process

Reliability methods used in qualitative research were also applied to the meta-thematic analysis in this study. During this phase, the concept of researcher triangulation (Streubert & Carpenter, 2011) was used, where two researchers collaborated throughout the meta-thematic analysis process. Furthermore, direct quotes from participant expressions were included to provide the raw data source while creating the themes and codes. The related quotes were expressed with codes indicating the study and page number from which the quotes were taken. For example, the numeric expression M12-s.15 refers to M (article), 12 (study number), and "s.15" (page number). The codes obtained from the YÖK National Thesis Center were also abbreviated using thesis numbers. Through the meta-analysis of studies conducted on primary school AI use, effect sizes were identified, and codes and themes were created through the meta-thematic analysis. To support these findings and provide a holistic table, the opinions of class teachers were analyzed using the Rasch measurement model.

2.3. Rasch Measurement Model Analysis Process

In the third dimension of the research, quantitative data related to teacher opinions on the use of AI in primary schools were analyzed using the Rasch measurement model, developed by Linacre (1993; 2008). This model is based on objectivity (Semerci, 2011). It also evaluates other variable sources that may affect the test results, such as item difficulty levels, raters, scoring keys, conditions, tasks, and scoring criteria (İlhan, 2016; Lynch & McNamara, 1998). The Multi-Faceted Rasch model establishes necessary linear relationships between different facets (such as AI application usage levels, evaluation item characteristics, and rater behaviors) and forms these connections (Hambleton & Swaminathan, 1985). In the current research, the Rasch model includes three facets: 21 primary school teachers, 18 items related to AI use in education, and four educational programs (Turkish, Mathematics, Life Science, and Science). To collect data, the researchers created the "Primary School Teachers' AI Application Usage Evaluation Form."

2.3.1. Study Group

The study group consisted of 21 classroom teachers working in state primary schools during the 2024-2025 academic year. The raters independently evaluated the teachers' AI application usage levels using the 18-item evaluation form.

2.3.2. Research Data and Analysis

To collect data for the study, the "Primary School Teachers' AI Application Usage Evaluation Form" was used (Appendix B). The evaluation form, developed within the framework of the Rasch measurement model and consisting of 18 items, assesses teachers' knowledge acquisition, knowledge deepening, and knowledge creation under the headings of curriculum and assessment, pedagogy, and the application of digital skills. The form includes evaluations of the Turkish, Mathematics, Science, and Life Science curricula. The evaluation form was created based on a literature review (Ministry of National Education [MEB], 2024; UNESCO, 2018) and expert opinions. Afterward, the form was re-examined by experts for verification, and necessary additions and deletions were made. This form contains a total of 18 criteria. The content validity index (CVI) of the items was calculated using the Content Validity Ratio (CVR) formula developed by Lawshe (1975) and found to be 0.81 (Appendix C). According to Veneziano and Hooper (1997), this value is statistically significant at the 0.05 level. The analysis of the AI application evaluation form was conducted using the FACETS analysis program developed by Linacre (1993: 2-15). FACETS is an analysis program frequently used in the multi-faceted Rasch measurement model. This program generally includes three main facets: rater, ability, and task (Güler & Gelbal, 2010).

3. Findings

This section interprets the findings obtained through meta-analysis, meta-thematic analysis, and participant opinions on various variables related to the use of AI applications in both classroom and out-of-class teaching environments. In the first phase of the study, meta-analytic findings based on document analysis are presented; in the second phase, results of the meta-thematic analysis; and in the third phase, participant opinions supported by quantitative data.

3.1. Meta-Analysis Findings on AI Applications

When examining the meta-analysis findings in Table 2, it was determined that the effect size of academic achievement scores (AA) for AI-supported applications, calculated using the REM method, was $g = .51$ [.28; .74]. This effect size is classified as moderate according to Thalheimer and Cook's (2002) classification, indicating that AI applications have a positive and beneficial effect on these variables. Additionally, a significant difference was found between the scores based on test types ($p < .05$).

Table 2. Meta-Analysis Data.

Test Type	Model	n	g	95 %+ Confidence Interval		Heterogeneity		
				Lower	Upper	Q	p	I ²
AA	FEM	24	0.59	0.47	0.64	163.11	0.00	85.90
	REM	24	0.51	0.28	0.74			

When examining the results of the heterogeneity test in Table 2, it is seen that the effect sizes of the scores for AI applications exhibit a heterogeneous distribution ($QB = 163.11$; $p < .05$). The I^2 value was calculated as 85.90%, indicating that the observed 86% variance originates from the true variance among the studies. According to Cooper et al. (2009), an I^2 value of 25% is considered low, around 50% is moderate, and 75% or higher is high heterogeneity. The calculated I^2 value of 85.90% in this study points to a high level of heterogeneity (Higgins et al., 2003). This suggests the presence of moderator variables affecting the total effect size. In other words, the detection of high heterogeneity indicates the need for a moderator analysis (Borenstein et al., 2009).

As presented in Table 3, the duration of implementation and sample size have been selected as moderator analyses. In the significance test, a significant difference was found in the duration of implementation ($QB = 7.69$; $p < .05$); however, no significant differences were observed in terms of

lessons (QB = 1.7; p > .05) and sample size (QB = 1.29; p > .05). Nevertheless, when the analysis results are generally evaluated, it can be stated that various variables have a moderate effect on artificial intelligence applications.

Table 3. General Effect Sizes of Studies Included in the Analysis According to Moderator Analysis.

Items	Groups	Effect Size and 95% Confidence Interval				Null Test		Heterogeneity		
		n	g	Lower Limit	Upper Limit	Z-value	P-value	Q-value	df	P-value
Application Duration	1-4	0.59	0.59	0.30	0.88	4.01	0.00			
	5+	0.09	0.09	-0.15	0.33	0.75	0.45			
	Unspecified	0.58	0.58	-0.02	1.19	1.89	0.06			
	Total	0.32	0.32	0.14	0.50	3.54	0.00	7.69	2	0.02
Subjects	Maths	19	0.44	0.18	0.71	3.25	0.01			
	AI	3	0.80	0.07	1.53	2.15	0.03			
	Others	2	0.81	0.19	1.44	2.54	0.01			
	Total	24	0.53	0.30	0.76	4.47	0.00	1.7	2	0.43
Sample Size	Small	6	0.50	0.09	0.90	2.40	0.02			
	Medium	9	0.37	0.18	0.55	3.93	0.00			
	Large	6	0.63	0.18	1.08	2.75	0.01			
	Total	24	0.42	0.26	0.57	5.24	0.00	1.29	2	0.52

3.2. Meta-Thematic Findings Regarding Artificial Intelligence Applications

In this part of the study, the themes and codes obtained through meta-thematic analysis of the studies accessed regarding the effects of various variables on artificial intelligence applications are presented. The data related to the effects of different variables on artificial intelligence applications were analyzed, and two themes were formed: the impact of artificial intelligence applications on educational settings and challenges encountered and solutions proposed. The themes and codes are illustrated in models. Additionally, direct quotations are provided as reference sentences within the comments.

When examining Figure 4, the contributions of artificial intelligence applications to learning environments can be seen in the model as codes. Some of the codes in this model can be explained as providing support for challenging subjects, presenting topics through visualization, enhancing problem-solving skills, fostering intrinsic and extrinsic motivation, supporting socio-emotional development, and improving presentation skills. Expressions and quotations that can serve as references to these codes are stated in the studies as follows:

“I like to learn English with it (the AI coach) as it helps improve my English competence,” and “You can get (virtual) flowers and awards if you practice English with the AI coach every day and achieve good performance” (M1-p. 6).

In another study: “He also started to improve his oral skills, and finally gave presentations in front of large audiences” (M2-p. 1890).

Additionally, in the study coded (M5-p. 12), it is stated as: “I thought it was interesting because I was able to actually control the AI to play rock-paper-scissors with only two fingers.”

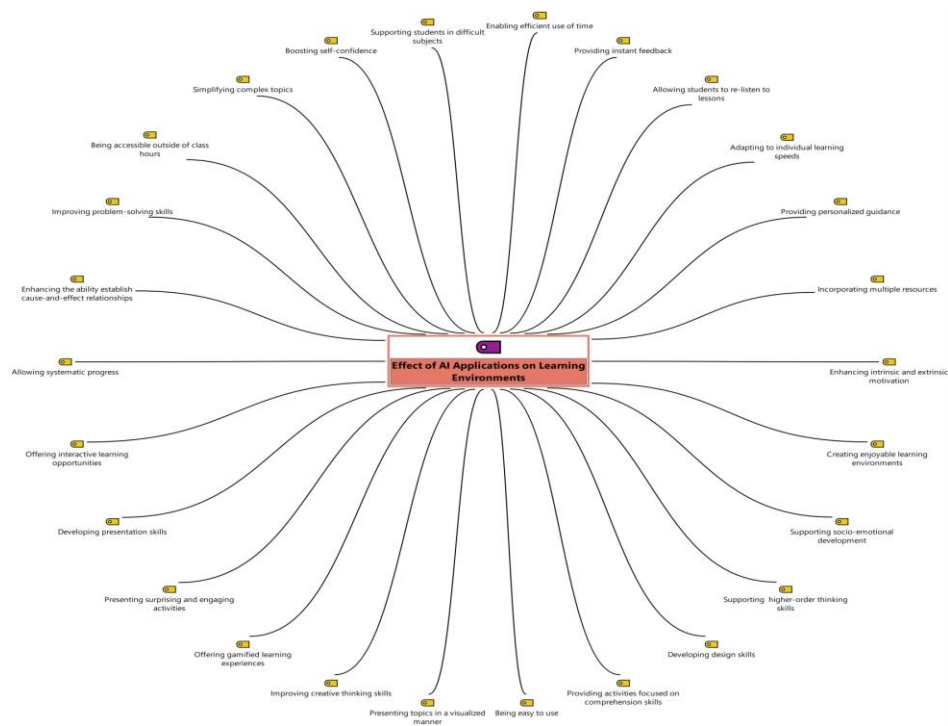


Figure 4. The Effect of Artificial Intelligence Applications on Learning Environments.

Furthermore, it can be stated that artificial intelligence applications help students develop critical thinking skills, promote learning through exploration, enhance problem-solving skills, and foster the acquisition of various other competencies. However, some challenges may be encountered during the implementation of these applications. The problems faced by practitioners and the proposed solutions are presented in the figure below.

When examining Figure 5, the problems encountered in learning environments during the artificial intelligence (AI) application process and the corresponding solutions are visualized. Some of these issues are indicated with the following codes: AI applications being deemed insufficient, difficulty in understanding questions asked, lack of knowledge about AI, inability to produce adequate outcomes, concerns about personal data being compromised, lack of emotional support from AI applications, and anxiety related to AI. Expressions related to these codes include:

“In my opinion, artificial intelligence is software that humans create, that we decide on the different things it learns and, after that, the computer adds them to other things it learned, like when we trained it with pictures, we also showed some pictures, and not always the computer was right, so we tried to give it more pictures and teach it things” (M3, p. 187).

In the study (M1, p. 10), the limitations of AI in sentence formation and vocabulary memorization were stated as: “For instance, one student indicated that ‘... it (the AI coach) improves our oral English through several ways, such as English shadowing, mimicking picture books, and memorizing vocabulary...’ Another student wrote, ‘it (the AI coach) contains many resources linked to our textbooks, picture books, and movie clips for budding practice.’”

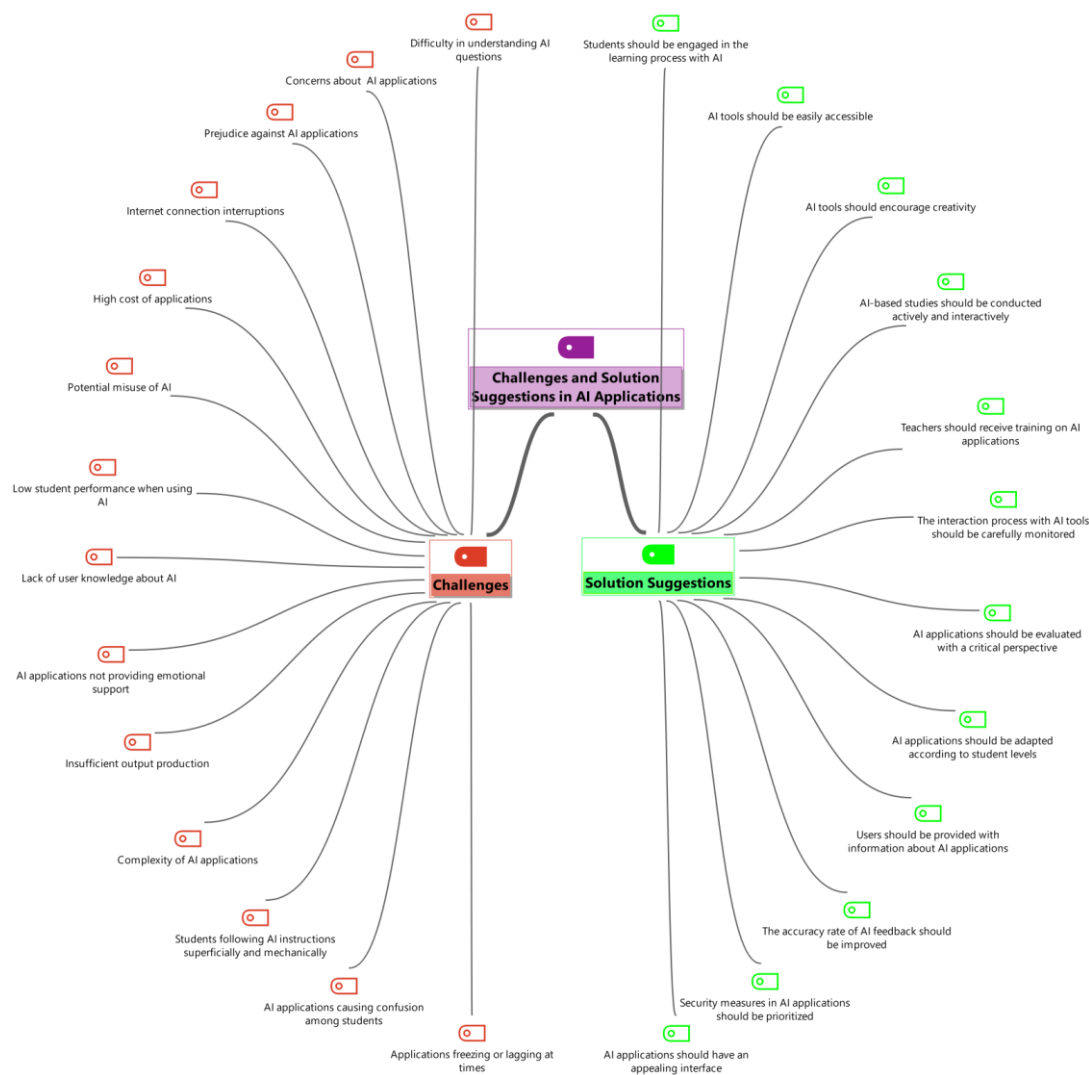


Figure 5. Problems Encountered in Artificial Intelligence Applications and Solution Suggestions.

Additionally, concerns about AI were expressed as follows: “I had concerns about the potential for AI to dominate the world, given its ability to complete tasks in just a few steps” (M5, p. 12).

During the analysis process, solution suggestions for potential problems in AI applications were identified. Some of the proposed solutions include: Students should actively engage in learning efforts with AI; AI applications should be used through teamwork; AI application activities should be conducted interactively and actively; security measures should be prioritized in AI applications.

Quotations referencing these suggestions include the following:

In the study (M2, p. 1822), “Students building basic robotic models benefit when they are working individually; meanwhile, students might work better in teams (ideally, two to three members per team) when working on advanced robotic models that include writing code (programming).”

Another example states: “In order to reduce such fraud, I think it is more convenient and accurate to use AI than humans” (M4, p. 12).

In conclusion, while certain issues may arise during AI applications, these problems can be resolved by utilizing the proposed solutions.

Findings derived from classroom teachers' opinions have been tabulated and expressed using the Rasch measurement model. This approach provided quantitative support to the data obtained through the mixed-meta process, presenting a detailed interaction of all the dynamics of the study.

3.4. Findings Related to the Rasch Measurement Model for Artificial Intelligence Applications

Since the research was conducted using a quantitatively supported mixed-meta method, this section presents the findings related to AI applications in primary school education using the multi-faceted Rasch measurement model. When analyzing the AI applications, the surfaces used in the study (teachers’ AI applications, judges’ strictness/generosity, and the appropriateness of the items used) and general information about these surfaces are provided in the calibration map shown in Figure 6.

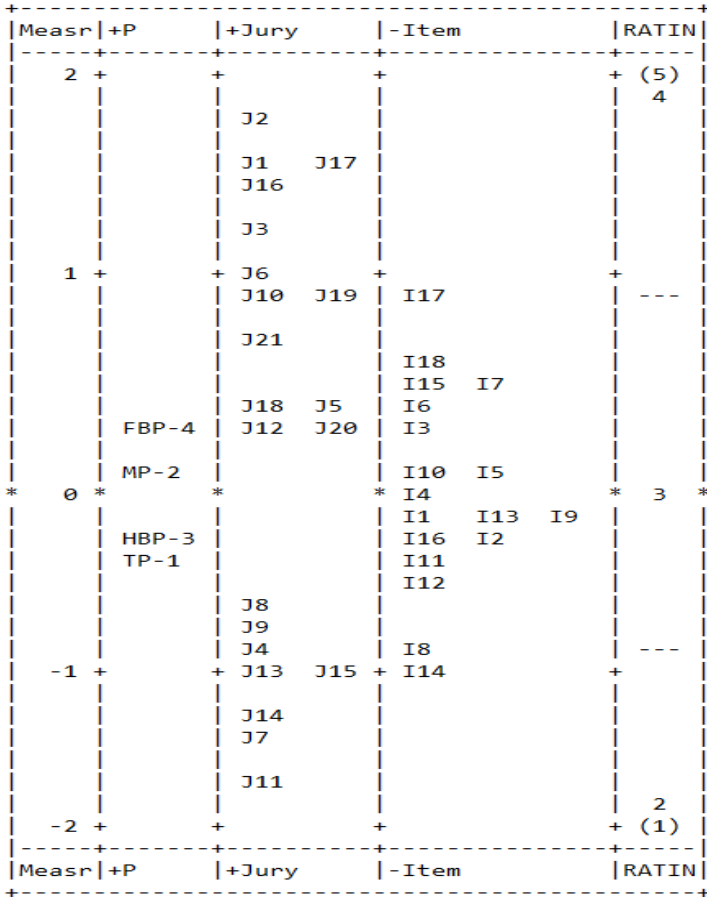


Figure 6. Data Calibration Map.

In the findings, four curricula (Turkish, mathematics, life sciences, and science education curricula), 21 judges, and 18 evaluation items related to the program content were taken into consideration. In this calibration map, it was found that the use of AI applications was at a high level in the science and mathematics curricula (Science Education Program: FBP-4 and Mathematics Program: MP-2), while it was observed to be at a low level in the life sciences and Turkish education programs (Life Sciences Program: HBP-3 and Turkish Program: TP-1).

Among the judges (raters), J2 was identified as the most generous rater, while J11 was determined to be the strictest rater.

When the column containing the items related to the use of AI applications was examined, the item “I can enable students to prepare a presentation describing their environment using AI tools” (I17) was found to be the most difficult item, whereas the item “I know how to efficiently use AI applications in classroom activities” (I14) was identified as the easiest item.

After evaluating the findings from the calibration map, the analysis report prepared for the curricula is presented in Figure 7.

Total Score	Total Count	Obsvd Average	Fair(M) Average	+ Measure	Model S.E.	Infit MnSq ZStd	Outfit MnSq ZStd	Estim. Discrm	Correlation PtMea PtExp	N	P
1233	378	3.26	3.28	.31	.07	.94 -.8	.95 -.7	1.07	.71 .65	4	FBP-4
1200	378	3.17	3.19	.14	.07	1.00 .0	1.00 .0	1.01	.71 .65	2	MP-2
1134	378	3.00	3.00	-.19	.07	1.09 1.3	1.09 1.2	.90	.59 .65	3	HBP-3
1119	378	2.96	2.96	-.27	.07	.94 -.9	.94 -.8	1.07	.59 .65	1	TP-1
1171.5	378.0	3.10	3.11	.00	.07	.99 -.1	1.00 -.1		.65	Mean (Count: 4)	
46.8	.0	.12	.13	.24	.00	.06 .9	.06 .9		.06	S.D. (Population)	
54.0	.0	.14	.15	.28	.00	.07 1.0	.07 1.0		.07	S.D. (Sample)	
Model, Populn: RMSE .07 Adj (True) S.D. .23 Separation 3.18 Strata 4.57 Reliability .91											
Model, Sample: RMSE .07 Adj (True) S.D. .27 Separation 3.72 Strata 5.29 Reliability .93											
Model, Fixed (all same) chi-square: 44.4 d.f.: 3 significance (probability): .00											
Model, Random (normal) chi-square: 2.8 d.f.: 2 significance (probability): .24											

Figure 7. Measurement Report of Curricula.

As a result of the Rasch analysis, the reliability coefficient was determined to be 0.91. This coefficient value indicates the reliability with which the curricula were ranked. Furthermore, when the data in Figure 7 were analyzed, it was observed that there were statistically significant differences between the curricula ($X^2=44.4$; $df=3$; $p=0.00$). Additionally, the standard error (RMSE) of the logit values related to the curricula was found to be 0.07, indicating a very low level of error.

This error rate, with its adjusted standard deviation, is below the critical value of 1.0, which is considered acceptable. In this context, the order of AI usage rates in the curricula from high-quality to low-quality can be listed as follows:

- FBP-4 (Science Education Curriculum)
- MP-2 (Mathematics Education Curriculum)
- HBP-3 (Life Sciences Curriculum)
- TP-1 (Turkish Education Curriculum)

In Rasch analysis, the quality control limits for "infit" (internal consistency) and "outfit" (outlier-sensitive consistency) values are accepted to be between 0.6 and 1.4 (Wright & Linacre, 1996). During the decision-making process, the "infit" values are used for unexpected responses from judges, while "outfit" values are used for unexpected distant responses (Baş Türk & Işıkoğlu, 2008; Batdı, 2014a). Upon examining the values in Figure 7, it is understood that these limits are appropriate.

In Figure 8, information regarding the strictness/leniency of the judges in relation to their use of AI in secondary education is presented. In Figure 7, it is observed that the judges' scores are ranked from the strictest to the most lenient. The judge identified with the code J2 is seen to be the most lenient, while the judge identified with the code J11 is the strictest. Additionally, the judge separation index was observed to be 6.44, and the reliability coefficient was noted as 0.98. A statistically significant difference was also found between the strictness/leniency of the judges' scores ($X^2=872.8$; $df=20$; $p=0.00$).

Total Score	Total Count	Obsvd Average	Fair(M) Average	+ Measure	Model S.E.	Infit MnSq ZStd	Outfit MnSq ZStd	Estim. Discrm	Correlation PtMea PtExp	Nu Jury
283	72	3.93	3.94	1.78	.17	.75 -1.6	.75 -1.6	1.30	.56 .35	2 J2
272	72	3.78	3.78	1.46	.17	.93 -.3	.94 -.3	1.04	.02 .36	1 J1
272	72	3.78	3.78	1.46	.17	.98 .0	1.00 .0	.99	.11 .36	17 J17
270	72	3.75	3.76	1.41	.17	1.13 .8	1.13 .8	.87	.61 .36	16 J16
261	72	3.63	3.63	1.16	.17	1.12 .7	1.11 .7	.87	.36 .36	3 J3
257	72	3.57	3.57	1.05	.16	.83 -1.0	.83 -1.0	1.18	.38 .36	6 J6
253	72	3.51	3.52	.94	.16	.95 -.2	.97 -.1	1.04	.41 .37	19 J19
252	72	3.50	3.50	.91	.16	1.00 .0	1.00 .0	.99	.09 .37	10 J10
243	72	3.38	3.38	.67	.16	1.23 1.3	1.24 1.4	.75	.29 .37	21 J21
233	72	3.24	3.24	.41	.16	1.28 1.6	1.28 1.6	.70	.55 .37	5 J5
232	72	3.22	3.22	.39	.16	1.29 1.7	1.29 1.6	.72	.25 .37	18 J18
228	72	3.17	3.17	.28	.16	.82 -1.1	.82 -1.1	1.20	.29 .37	12 J12
228	72	3.17	3.17	.28	.16	.33 -5.7	.33 -5.7	1.70	.22 .37	20 J20
190	72	2.64	2.64	-.69	.16	.96 -.2	.96 -.1	1.05	.51 .37	8 J8
187	72	2.60	2.60	-.77	.16	1.10 .6	1.11 .6	.90	.56 .37	9 J9
181	72	2.51	2.51	-.92	.16	.78 -1.4	.78 -1.4	1.26	.62 .37	4 J4
179	72	2.49	2.48	-.98	.16	.81 -1.2	.82 -1.1	1.21	.38 .37	13 J13
178	72	2.47	2.47	-1.00	.16	1.48 2.6	1.48 2.6	.47	.30 .37	15 J15
170	72	2.36	2.36	-1.21	.16	.98 .0	1.00 .0	1.01	.14 .37	14 J14
167	72	2.32	2.32	-1.29	.16	.93 -.4	.92 -.4	1.10	.62 .37	7 J7
150	72	2.08	2.08	-1.76	.17	1.16 1.0	1.14 .9	.81	.16 .36	11 J11
223.1	72.0	3.10	3.10	.17	.16	.99 -1.1	1.00 -1.1		.35	Mean (Count: 21)
40.8	.0	.57	.57	1.07	.00	.24 1.7	.24 1.7		.19	S.D. (Population)
41.8	.0	.58	.58	1.10	.00	.24 1.7	.24 1.7		.19	S.D. (Sample)

Model, Populn: RMSE .16 Adj (True) S.D. 1.06 Separation 6.44 Strata 8.92 Reliability .98
Model, Sample: RMSE .16 Adj (True) S.D. 1.08 Separation 6.60 Strata 9.14 Reliability .98
Model, Fixed (all same) chi-square: 872.8 d.f.: 20 significance (probability): .00
Model, Random (normal) chi-square: 19.6 d.f.: 19 significance (probability): .42

Figure 8. the strictness/leniency of the judges.

As a result, it was determined that the score observed in J2, with 283, was the most lenient, while the score observed in J11, with 150, was the strictest.

When the infit (internal consistency) and outfit (outlier-sensitive consistency) values of the surfaces in Figure 8 were evaluated, it was found that the judge identified with the code J15 did not meet the infit and outfit criteria (range of 0.6 to 1.4). In this case, it can be stated that the mean of the infit and outfit squares for the judge coded as J15 fell outside the defined limits. This can be interpreted as the lenient judge (J15) not exhibiting consistent scoring behavior while evaluating AI applications.

On the other hand, it was noted that the other judges (17 judges) demonstrated consistency in scoring among themselves, as their values fell within the expected quality control range, and thus, they can be considered as appropriate.

Upon examining the results of the data analysis presented in Figure 9, it was observed that the separation index is 3.04, and the reliability coefficient is 0.90. This reliability coefficient indicates that the items used in the study to determine teachers' levels of using AI tools are reliable. Additionally, there were statistically significant differences among the item difficulty levels used to evaluate teachers' opinions regarding their use of AI tools ($X^2=181.7$; $df=17$; $p=0.00$).

Total Score	Total Count	Obsvd Average	Fair(M) Average	- Measure	Model S.E.	Infit MnSq	ZStd	Outfit MnSq	ZStd	Estim. Discrm	Correlation PtMea	PtExp	Nu Item
221	84	2.63	2.62	.90	.15	1.03	.2	1.04	.2	.99	.74	.63	17 I17
235	84	2.80	2.79	.58	.15	1.16	1.1	1.17	1.1	.83	.77	.63	18 I18
238	84	2.83	2.83	.51	.15	.83	-1.1	.83	-1.1	1.21	.68	.63	7 I7
238	84	2.83	2.83	.51	.15	.66	-2.5	.66	-2.5	1.37	.66	.63	15 I15
243	84	2.89	2.89	.40	.15	.94	-.3	.94	-.3	1.05	.66	.63	6 I6
246	84	2.93	2.93	.33	.15	.99	.0	1.00	.0	1.00	.49	.63	3 I3
254	84	3.02	3.03	.15	.15	1.01	.1	1.01	.1	.98	.73	.63	5 I5
254	84	3.02	3.03	.15	.15	.95	-.3	.95	-.2	1.05	.70	.63	10 I10
259	84	3.08	3.09	.03	.15	.73	-1.9	.73	-1.9	1.29	.67	.63	4 I4
266	84	3.17	3.18	-.13	.15	1.13	.8	1.15	.9	.85	.57	.63	1 I1
267	84	3.18	3.19	-.15	.15	.81	-1.3	.81	-1.3	1.22	.74	.63	9 I9
267	84	3.18	3.19	-.15	.15	1.07	.5	1.06	.4	.93	.54	.63	13 I13
268	84	3.19	3.20	-.17	.15	.76	-1.7	.76	-1.6	1.25	.61	.63	2 I2
271	84	3.23	3.24	-.24	.15	1.51	3.0	1.52	3.0	.42	.51	.63	16 I16
273	84	3.25	3.26	-.29	.15	.79	-1.4	.78	-1.5	1.24	.66	.63	11 I11
285	84	3.39	3.41	-.56	.15	1.27	1.7	1.25	1.5	.72	.42	.63	12 I12
298	84	3.55	3.57	-.87	.15	1.22	1.4	1.24	1.5	.73	.42	.63	8 I8
303	84	3.61	3.63	-.99	.16	1.02	.2	1.01	.1	1.00	.72	.62	14 I14
260.3	84.0	3.10	3.11	.00	.15	.99	-.1	1.00	-.1		.63		Mean (Count: 18)
21.2	.0	.25	.26	.49	.00	.21	1.4	.21	1.4		.11		S.D. (Population)
21.8	.0	.26	.27	.50	.00	.22	1.4	.22	1.4		.11		S.D. (Sample)
Model, Populn: RMSE .15 Adj (True) S.D. .46 Separation 3.04 Strata 4.39 Reliability .90													
Model, Sample: RMSE .15 Adj (True) S.D. .48 Separation 3.14 Strata 4.52 Reliability .91													
Model, Fixed (all same) chi-square: 181.7 d.f.: 17 significance (probability): .00													
Model, Random (normal) chi-square: 15.6 d.f.: 16 significance (probability): .48													

Figure 9. Item Difficulty Analyzes for Evaluating Artificial Intelligence Applications.

The most difficult items related to the use of AI applications were identified as:

“I can guide students to prepare a presentation describing their environment using AI tools.”

“I can enable students to create simple patterns (e.g., drawing, storytelling) using AI tools.”

“I can integrate AI applications into pedagogical methods and techniques to contribute to students’ learning processes.”

On the other hand, the easiest items were:

“I know how to use AI applications effectively in classroom activities.”

“I know how AI applications can support teaching activities in a classroom setting.”

“I can encourage students to collaborate and improve their skills using AI tools.”

The standard error (RMSE) for the analysis of criteria prepared by the researchers to determine the level of teachers' use of AI applications was found to be 0.15, indicating that the standard error related to identifying the levels of AI application use is quite low.

Furthermore, when the infit (internal consistency) and outfit (outlier-sensitive consistency) statistical values for the research surfaces in Figure 9 were examined, it was noted that only the item coded as I16 exceeded the outfit threshold. In this context, it can be stated that the mentioned item demonstrates inconsistency in evaluating the levels of AI use.

All other items, however, were within the specified consistency value limits, indicating their reliability in assessing AI application usage levels.

4. Discussion and Conclusions

In this study, a mixed-meta method, integrated with a quantitative research process, was employed to examine the effects of studies conducted at the primary school level on various variables and to reveal the current state of the literature in comparison with other studies.

In line with the purpose of the research, the following methodological steps were followed:

Meta-analysis was conducted first.

This was followed by meta-thematic analysis.

Finally, the participants' opinions were collected in the quantitative dimension and evaluated using the Rasch measurement model.

4.1. Results of the Meta-Analysis Process

The study includes 37 quantitative studies for the meta-analysis, while 4 qualitative studies were included in the meta-thematic analysis process. In the quantitative scope, the teachers' views on the relevant variables were collected using the prepared evaluation form. The findings based on the data obtained from the analysis are presented in this section. According to the meta-analysis, the results of the included studies based on the REM showed that the effect of various variables on the use of artificial intelligence was moderate ($g = .51$ [.28; .74]), based on Thalheimer and Cook's (2002) classification. The results of other studies in the literature, which were not included in this study, support the research results (Gilad & Levin, 2021; Kajiwarra et al., 2023). Based on these results, it can be stated that different variables have an impact on the use of artificial intelligence applications.

During the analysis, a moderator analysis was also conducted regarding the effects of various variables on the use of artificial intelligence. The first moderator analysis, related to the duration of implementation, indicated that a duration of 1-4 weeks had an effect size ($g = 0.59$). Secondly, when subjects were used as moderators, the highest effect size ($g = 0.81$) was found for other subjects. Thirdly, regarding sample size as a moderator, large sample groups showed the highest effect size ($g = 0.63$). When analyzing the effect sizes of moderators, a significant difference was observed for the duration of implementation ($QB = 7.69$; $p < .05$), while no significant difference was found for subjects ($QB = 1.7$; $p > .05$) and sample size ($QB = 1.29$; $p > .05$). The findings of studies in the literature, such as those by Kablan et al. (2013) and Camnalbur et al. (2018), support the findings of this study in terms of subjects and sample size. In contrast to the studies in the literature, no effect on academic achievement was observed with time durations, but this research found a significant difference in the moderator related to implementation time. When considering the results of the moderator analyses as a whole, it can be stated that artificial intelligence applications have a moderate impact across all groups with similar effect sizes. During the meta-analysis process, it was observed that the majority of studies regarding artificial intelligence applications were conducted in mathematics classes. In this context, further research on the use of artificial intelligence applications in other subjects is recommended.

4.2. Results of the Meta-Thematic Analysis Process

In the second phase, a document analysis-based meta-thematic analysis was conducted, integrating the findings from the meta-analysis to validate them and expand the scope of the results. The themes and codes regarding the effects of artificial intelligence applications on learning environments and the problems and solutions encountered during their use were formed. The analysis revealed that students were in an interactive learning environment while using artificial intelligence applications. It can be stated that being highly interactive in the constructivist learning process helps make artificial intelligence applications more comprehensible (Shamir & Levin, 2021). One of the codes reached in the research was that artificial intelligence applications support individuals who learn at different paces and paths. Similar findings were found in Pilco (2020). It can be said that an individual's self-confidence in any task contributes to their success in that task. Our research concluded that artificial intelligence applications increase individuals' self-confidence and enhance their creativity. This result is supported by findings from other studies using digital technology to support student learning (Bers et al., 2014; Fitton et al., 2013).

In our research, it was also found that artificial intelligence applications improve problem-solving skills, creative thinking, and higher-order thinking abilities, aligning with findings in the literature indicating that technology-supported lesson content enhances students' cognitive capacities (Garrison, 2017; Ke, 2010).

Another theme created in our study is the problems and solutions encountered during the use of artificial intelligence applications. Studies indicating that individuals with low self-efficacy experience high computer anxiety (Ellis & Allaire, 1999) align with the findings in this research. The research findings show that students' anxiety while using artificial intelligence applications may be related to their low self-efficacy in this area. One of the codes found was that students were disappointed by the responses provided by artificial intelligence during the process. Similar to this

finding, the research of Toivonen et al. (2020) suggests that poor performance of the applications does not meet students' expectations. Another issue identified in the study is that artificial intelligence applications do not provide emotional support. Literature review shows that lack of emotional support in artificial intelligence applications can negatively affect the targeted lesson outcomes (Randall, 2019; Wang, 2024). Another identified issue is that students' feedback on artificial intelligence usage is not understood by the application. Findings from studies in the literature (Chen et al., 2020; Wang et al., 2018) support this finding in our research.

It is suggested that artificial intelligence applications should have an attractive design, as this will encourage students to develop a positive attitude towards using them (Reeves et al., 2020). Based on this, it can be concluded that artificial intelligence applications should be produced with engaging designs. Another problem encountered in the process is that educators do not have the necessary training, which prevents them from obtaining the desired results from the application. Our research found that teachers need sufficient knowledge and skills regarding artificial intelligence applications to overcome this issue. Pilco's (2020) study also aligns with our findings, indicating that teachers should receive technical training on artificial intelligence.

4.3. Results Related to the Rasch Measurement Model Process

The quantitative section of our research presents the findings related to teachers' views on their level of use of artificial intelligence applications, derived through analyses using the Rasch measurement model. In this way, the findings obtained through the mixed-meta processes were supported quantitatively to ensure the alignment of the study's results. The Rasch measurement model adopts a fundamental approach that relates the probability of answering a question correctly to an individual's ability (Baştürk, 2010). The multi-faceted Rasch measurement model, developed by John M. Linacre, not only examines the relationship between individuals' ability levels and the difficulty levels of the items on the measurement tool but also allows the evaluation of other variable sources that might affect test results, such as scorers, scoring keys, conditions, tasks, and scoring criteria (Lynch & McNamara, 1998). In this context, simultaneous surfaces (the level of use of artificial intelligence applications, the rigidity/generosity of the jury, and the characteristics of the evaluation questions) that were prioritized and analyzed in the multi-faceted Rasch measurement model have been ranked among themselves. Awareness of artificial intelligence applications, a concept related to Industry 4.0, is crucial as these applications play a significant role in shaping the future of technology (Doğan & Baloğlu, 2020). In the current study, teachers' views were examined in the context of curricula by associating dimensions such as curriculum and assessment, pedagogy, and the application of digital skills with sub-dimensions like knowledge acquisition, knowledge creation, and knowledge deepening.

The analysis revealed that the science and mathematics curricula had the highest level of use in the teaching programs. Artificial intelligence, which has made significant advancements in the past 50 years, has become an important research area (Talan, 2021). AI encompasses many cognitive areas of human intelligence, including learning, reasoning, planning, problem-solving, perception, natural language processing, deep learning, expert systems, image processing, sentiment analysis, speech recognition, and more (Goodfellow et al., 2016). Therefore, it can be said that the primary use of AI in science and mathematics is due to the functionalities of artificial intelligence itself.

When the scorers' rigidity/generosity information was evaluated in relation to the assessment of AI applications, it was found that the jury defined by the J2 code was the most generous, while the jury defined by the J11 code was the strictest. Furthermore, it was determined that the scorers were reliably ranked in terms of rigidity/generosity and differed from each other. Studies using the multi-faceted Rasch measurement model in the literature also indicate that scorers (juries) can be both objective and biased at times (Baştürk, 2010; Batdı, 2014a; Köse et al., 2016; Semerci, 2012). In addition, the analysis concluded that the jury separation index was 6.44, which is above the desired level. This value indicates that there are differences in scoring among the juries, that juries vary according to

their generosity/rigidity levels, and that scoring errors related to generosity/rigidity exist in the scores given by the juries (Uluman & Tavşancıl, 2017).

The items in the teacher evaluation form regarding the level of use of AI applications were found to serve the purpose of measuring teachers' competence levels. When the items prepared to determine teachers' level of use of AI were analyzed, it was found that the most difficult items were "I can make students prepare a presentation describing their environment using AI tools," "I can make students prepare simple designs (like drawing pictures, creating stories) using AI tools," and "I can integrate AI applications into pedagogical methods and techniques to support students' learning processes." Therefore, it can be said that teachers may face difficulties in using AI. Based on these results, it can be expressed that it is an area that needs further research on how AI can be used effectively in educational environments to achieve the desired outcomes and create designs based on the characteristics of the gains. On the other hand, the easiest items were "I know how to use AI applications effectively in classroom activities," "I know how AI applications can support teaching activities in the classroom," and "I can encourage students to collaborate and develop using AI tools." Although the participants possess basic knowledge about AI, it can be said that teachers need to receive necessary training to stay up-to-date with developing technologies (Benvenuti et al., 2023).

When considering the results of the entire research process as a whole, it is evident that the obtained data support each other. In the meta-analysis process, it was found that AI studies in elementary schools were most frequently conducted in mathematics lessons (Zhang & Aslan, 2021). In the Rasch measurement model section of the study, it was observed that classroom teachers most commonly used AI applications in science and mathematics lessons. In this regard, it can be said that the results of these two processes overlap. Furthermore, during the meta-thematic analysis process, the theme of training practitioners in AI applications emerged as an issue and solution. In the Rasch measurement model in the jury opinions section, it was found that the most difficult items were related to teachers preparing presentations with AI, integrating AI applications into the curriculum, and preparing simple designs related to teaching practices, indicating that educators need training in AI applications. Thus, the results of the study's processes are compatible. Moreover, during the meta-thematic analysis, the theme related to the effect of AI on educational environments was "providing interactive teaching opportunities," which aligns with the item in the Rasch measurement model "I can encourage students to collaborate and develop using AI tools" being among the easiest items. In this way, it can be said that the results from the two processes are integrated, as both contribute to the creation of an interactive classroom environment where teachers facilitate collaboration among students using AI tools (Lin et al., 2023). In today's technological age, the findings of this study, which involves both qualitative and quantitative research within the methodological pluralism framework, reflect how artificial intelligence, as the most prominent tool of the connectivity paradigm in modern educational environments, encapsulates and supports many of the findings.

4.4. Limitations

Although this study presents a mixed-methods approach supported by quantitative analysis within the framework of methodological pluralism, there are several limitations to consider. The data collected for the meta-analysis and meta-thematic analysis were limited to certain databases. The meta-analysis focused solely on the impact of AI applications on academic achievement. Moderator variables in the study included the duration of the intervention, the subject areas, and the sample size. In the Rasch measurement model, data was collected regarding the Turkish, mathematics, science, and life sciences curricula. Additionally, survey questions were developed based on teachers' digital competencies (UNESCO, 2018), covering aspects such as curriculum and assessment, pedagogy, and digital skills. Finally, the study focused on gathering perspectives from classroom teachers.

4.5. Recommendations

- The application duration, subject areas, and sample sizes in AI-related research have significant effects on academic success and the impact of AI on educational environments. The use of the mixed-meta method, supported by the Rasch measurement model, has provided a more holistic perspective, allowing for a deeper exploration of the topic. Based on the limitations and findings of the study, the following recommendations are made:
- Research on AI applications in primary school subject areas such as art, music, and physical education can be conducted. In addition to quantitative methods, qualitative methods could be employed to explore the effectiveness and applicability of survey questions.
- The meta-analysis phase of the study could include investigations into the impact of AI applications on attitudes and long-term retention.
- Studies could explore teachers' information and technology competencies (UNESCO, 2018) within other professional practice areas.
- The study focused on perspectives from classroom teachers. Including evaluators from different expertise levels could broaden the scope of the study.
- Despite teachers' positive expectations regarding AI, it is essential that they first familiarize themselves with the technology and learn how to integrate it into their classrooms. Many teachers may regard AI as an advanced technological product without prior experience. In this regard, in-service training could increase teachers' knowledge about AI and improve their integration of this technology, significantly enhancing student success and the learning experience (Kim NJ and Kim MK, 2022).
- Given the methodological diversity, the use of a mixed-meta method combined with quantitative analysis has allowed for a comprehensive examination of the findings, with detailed insights into how various variables affect the use of AI applications. Therefore, it is recommended to apply the mixed-meta method integrated with either qualitative or quantitative analyses in other areas to achieve comprehensive research findings.
- Policymakers should take necessary measures to address concerns related to ethics, data security, and human rights as AI becomes more integrated into education.

Appendix A. Agreement Value Ranges of Themes Related to Artificial Intelligence Applications

Effect on Learning Environments					Problems Encountered				Related Solution Suggestions				Problems Encountered and Solution Suggestions						
K2					K2				K2				K2						
Σ	+	-	Σ		Σ	+	-	Σ		Σ	+	-	Σ		Σ	+	-	Σ	
	+	26	2	28		+	14	1	15		+	12	1	13		+	26	2	28
	-	3	18	21		-	0	9	9		-	1	7	8		-	1	16	17
	Σ	29	20	49		Σ	14	10	24		Σ	13	8	21		Σ	27	18	45
Kappa: .790 p: .000					Kappa: .913 p: .000				Kappa: .798 p: .000				Kappa: .860 p: .000						

Appendix B. Primary School Teachers' Artificial Intelligence Applications Evaluation Form

Değerli öğretmenimiz, bu uygulamanın amacı “İlkokul Öğretmenlerinin Yapay Zeka Uygulamalarını Kullanma Düzeylerini Belirlenen Kriterler Çerçevesinde Değerlendirmektir.” Lütfen maddelerin karşısında yer alan her seçenekten size göre en uygun olanını, altındaki kutucuğa “5: Tamamen 4: Çoğunlukla 3: Kısmen 2: Nadiren 1: Hiç” şeklinde rakamla belirtiniz.							
Alt Boyutlar	Madde No	1. Cinsiyetiniz? ☐ Erkek ☐ Kadın	Öğretim Programları				
			Türkçe	Mat.	Hay. Bil.	Fen Bil.	
Müfredat ve Değerlendirme	Bilgi Edinimi	1	Ders içeriklerimi desteklemek için yapay zeka tabanlı kaynakları müfredata entegre edebilirim.				
		2	Öğrencilerin ilerlemesini değerlendirmek için yapay zeka uygulamalarından yararlanırım.				
	Bilgi Derinleş.	3	Öğretim programını yapay zeka uygulamalarını kullanarak güncelleyip zenginleştirebilirim.				
		4	Öğrencilerin dijital ortamda gösterdikleri performansı değerlendirirken yapay zeka araçlarını kullanabilirim.				
	Bilgi Oluş.	5	Yapay zeka araçları ile ders değerlendirmesi yaparken öğrencilerin bireysel öğrenme düzeylerini daha iyi analiz edebilirim.				
		6	Yapay zeka tabanlı değerlendirme araçlarını kullanarak öğrencilere bireysel geribildirim sağlayabilirim.				
Pedagoji	Bilgi Edinimi	7	Öğrencilerin öğrenme süreçlerine katkı sağlamak için yapay zeka uygulamalarını pedagojik yöntem ve teknikler entegre edebilirim.				
		8	Yapay zeka uygulamalarının sınıf ortamındaki öğretim etkinliklerini ne şekilde destekleyebileceğini bilirim.				
	Bilgi Derinleş.	9	Ders planlaması sırasında yapay zeka araçlarını etkili bir şekilde kullanarak öğrencilere daha zengin öğrenme deneyimleri sunabilirim.				
		10	Öğrencilerin konuyu daha iyi kavramaları için bilgiye dayalı yapay zeka destekli dijital içerikler (örneğin etkileşimli ders notları, videolar) oluşturabilirim.				
	Bilgi Oluş.	11	Öğrencilerin aktif katılımını sağlamak amacıyla yapay zeka araçlarını kullanabilirim.				
		12	Yapay zeka araçlarıyla öğrencilerin iş birliği yapmalarını ve geliştirmelerini teşvik edebilirim.				
Dijital Becerilerin Uygulanması	Bilgi Edinimi	13	Öğrencilerin yapay zeka araçlarıyla proje yapmalarına (resim yapma, hikaye oluşturma gibi) rehberlik edebilirim.				
		14	Yapay zeka uygulamalarını sınıf içi etkinliklerde nasıl verimli kullanabileceğimi bilirim.				
	Bilgi Derinleş.	15	Yapay zeka araçlarını kullanarak öğrencilerin iş birliği içinde çalışmalarını sağlayabilirim.				
		16	Öğrencilerin araştırma yaparken dijital araçları güvenli ve etkili bir şekilde kullanmalarını sağlayabilirim.				
	Bilgi Oluş.	17	Öğrencilerin yapay zeka araçlarıyla kendi çevrelerini anlatan bir sunu hazırlamalarını sağlayabilirim.				
		18	Öğrencilerin yapay zeka araçlarını kullanarak (resim yapma, hikaye oluşturma gibi) basit desenler hazırlamalarını sağlayabilirim.				

Appendix C. Content Validity Rates of Items for Evaluation of Artificial Intelligence Applications

MN	Maddeler	Gerekli	Yararlı/ Yetersiz	Gereksiz	KGO*
1	Ders içeriklerini desteklemek için yapay zeka tabanlı kaynakları müfredata entegre edebilirim.	12	-	-	%100
2	Öğrencilerin ilerlemesini değerlendirmek için yapay zeka uygulamalarından yararlanırım.	12	-	-	%100
3	Öğretim programını yapay zeka uygulamaları kullanarak güncelleyip zenginleştirebilirim.	11	1	-	%83
4	Öğrencilerin dijital ortamda gösterdikleri performansı değerlendirirken yapay zeka araçlarını kullanabilirim.	12	-	-	%100
5	Yapay zeka araçları ile ders değerlendirmesi yaparken öğrencilerin bireysel öğrenme düzeylerini daha iyi analiz edebilirim.	10	1	1	%67
6	Yapay zeka tabanlı değerlendirmeye araçlarını kullanarak öğrencilere bireysel geribildirim sağlayabilirim.	11	1	-	%83
7	Öğrencilerin öğrenme süreçlerine katkı sağlamak için yapay zeka uygulamalarını pedagojik yöntem ve teknikler entegre edebilirim.	10	1	1	%67
8	Yapay zeka uygulamalarının sınıf ortamındaki öğretim etkinliklerini ne şekilde destekleyebileceğini bilirim.	12	-	-	%100
9	Ders planlaması sırasında yapay zeka araçlarını etkili bir şekilde kullanarak öğrencilere daha zengin öğrenme deneyimleri sunabilirim.	11	--	1	%83
10	Öğrencilerin konuyu daha iyi kavramaları için bilgiye dayalı yapay zeka destekli dijital içerikler (örneğin etkileşimli ders notları, videolar) oluşturabilirim.	10	-	2	%67
11	Öğrencilerin aktif katılımını sağlamak amacıyla yapay zeka araçlarını kullanabilirim.	12	-	-	%100
12	Yapay zeka araçlarıyla öğrencilerin iş birliği yapmalarını ve geliştirmelerini teşvik edebilirim.	10	2	-	%67
13	Öğrencilerin yapay zeka araçlarıyla proje yapmalarına (resim yapma, hikaye oluşturma gibi) rehberlik edebilirim.	10	1	1	%67
14	Yapay zeka uygulamalarını sınıf içi etkinliklerde nasıl verimli kullanacağımı bilirim.	10	1	1	%67
15	Yapay zeka araçlarını kullanarak öğrencilerin iş birliği içinde çalışmalarını sağlayabilirim.	10	-	2	%67
16	Öğrencilerin araştırma yaparken dijital araçları güvenli ve etkili bir şekilde kullanmalarını sağlayabilirim.	12	-	-	%100
17	Öğrencilerin yapay zeka araçlarıyla kendi çevrelerini anlatan bir sunu hazırlamalarını sağlayabilirim.	10	-	2	%67
18	Öğrencilerin yapay zeka araçlarını kullanarak (resim yapma, hikaye oluşturma gibi) basit desenler hazırlamalarını sağlayabilirim.	10	1	1	%67

*Uzman Sayısı:12, Kapsam Geçerlik Ölçütü:0.56
Kapsam Geçerlik & İndeksi:%81 KGI > KGO (81> 0.56)

References

1. Akerkar, R. (2014). *Introduction to artificial intelligence*. PHI Learning Press.
2. Anıl, Ö. ve Batdı, V. (2023). Use of augmented reality in science education: A mixed-methods research with the multi-complementary approach. *Education and Information Technologies*, 28(5), 5147-5185. <https://doi.org/10.1007/s10639-022-11398-6>
3. Avci, O., Abdeljaber, O., Kiranyaz, S., Hussein, M., Gabbouj, M. ve Inman, D. J. (2021). A review of vibration-based damage detection in civil structures: From traditional methods to machine learning and deep learning applications. *Mechanical Systems and Signal Processing*, 147, Article 107077. <https://doi.org/10.1016/j.ymssp.2020.107077>
4. Baştürk, R. (2010). Bilimsel araştırma ödevlerinin çok yüzeyli Rasch ölçme modeli ile değerlendirilmesi. *Journal of Measurement and Evaluation in Education and Psychology*, 1(1), 51-57.
5. Baştürk, R. ve Işıkoğlu, N. (2008). Okul öncesi eğitim kurumlarının işlevsel kalitelerinin çok yüzeyli Rasch modeli ile analizi. *Kuram ve Uygulamada Eğitim Bilimleri*, 8(1), 7-32.
6. Batdı, V. (2014a). Ortaöğretim matematik öğretim programı içeriğinin Rasch ölçme modeli ve Nvivo ile analizi. *Turkish Studies*, 9(11), 93-109. <http://dx.doi.org/10.7827/TurkishStudies.7074>
7. Batdı, V. (2014b.). Kavram haritası tekniği ile geleneksel öğrenme yönteminin kullanılmasının öğrencilerin başarıları, bilgilerinin kalıcılığı ve tutumlarına etkisi: bir meta-analiz çalışması, *Dumlupınar Üniversitesi Sosyal Bilimler Dergisi*, 42, 93-102

8. Batdı, V. (2019). Meta-thematic analysis. In V. Batdı (Ed.), *Meta-thematic analysis: sample applications* (pp. 10–76). Anı Publication.
9. Batdı, V. (2021). Yabancılar dil öğretiminde teknolojinin kullanımı: bir karma-meta yöntemi. *Milli Eğitim Dergisi*, 50(1), 1213-1244. <https://doi.org/10.37669/milliegitim.942631>
10. Batdı, V. (2024a). Farklı değişkenlerin öznel iyi oluş düzeyine etkisi: nitel analizle bütünleştirilmiş karma-meta yöntemi. *Gaziantep Üniversitesi Eğitim Bilimleri Dergisi*, 8(1), 53-83.
11. Batdı, V. (Ed.). (2024b). *Meta-tematik analiz örnek uygulamalar*. Anı Yayıncılık.
12. Benvenuti, M., Cangelosi, A., Weinberger, A., Mazzoni, E., Benassi, M., Barbaresi, M. ve Orsoni, M. (2023). Artificial intelligence and human behavioral development: A perspective on new skills and competences acquisition for the educational context. *Computers in Human Behavior*, 148. <https://doi.org/10.1016/j.chb.2023.107903>
13. Barnard, Y. F. ve Sandberg, J. A. (1988). Applying artificial intelligence insights in a CAI program for “open sentence” mathematical problems in primary schools. *Instructional Science*, 17(3), 263-276. <https://doi.org/10.1007/BF00048345>
14. *Bers, M. U., Flannery, L., Kazakoff, E. R. ve Sullivan, A. (2014). Computational thinking and tinkering: Exploration of an early childhood robotics curriculum. *Computers & education*, 72, 145-157. <https://doi.org/10.1016/j.compedu.2013.10.020>
15. Borenstein, M., Hedges, L. V., Higgins, J. P. T. ve Rothstein, H. R. (2009). *Introduction to Meta-analysis*. Wiley.
16. Bowen, G.A. (2008). Naturalistic inquiry and the saturation concept: A research note. *Qualitative Research*, 8(1), 137–152. <https://doi.org/10.1177/1468794107085301>
17. Braccaccio, R., Hojaij, F. ve Notargiacomo, P. (2019). Gamification in the study of anatomy: The use of artificial intelligence to improve learning. *The FASEB Journal*, 33(1_supplement), 444-28. DOI: 10.1096/fasebj.2019.33.1_supplement.444.28
18. Bryman, A. (2016). *Social research methods*. Oxford university press.
19. *Bush, J. B. (2021). Software-based intervention with digital manipulatives to support student conceptual understandings of fractions. *British Journal of Educational Technology*, 52(6), 2299-2318. <http://dx.doi.org/10.1111/bjet.13139>
20. Camnalbur, M. (2008). *Bilgisayar destekli öğretimin etkililiği üzerine bir meta analiz çalışması* [Yayımlanmamış yüksek lisans tezi]. Marmara Üniversitesi.
21. Candan, F. ve Başaran, M. (2023). A meta-thematic analysis of using technology-mediated gamification tools in the learning process. *Interactive Learning Environments*, 1–17. <https://doi.org/10.1080/10494820.2023.2172589>
22. Chen, L., Chen, P. ve Lin, Z. (2020). Artificial intelligence in education: a review. *IEEE Access*, 8, 75264-75278. <https://doi.org/10.1109/access.2020.2988510>
23. Chen, S-Y., Lin, P-H. ve Chien, W-C. (2022). Children’s digital art ability training system based on ai-assisted learning: a case study of drawing color perception. *Front in Psychol.* 13. doi: 10.3389/fpsyg.2022.823078
24. Chiu, T. K. (2021). A Holistic approach to the design of Artificial Intelligence (AI) education for k-12 schools. *TechTrends*, 1–12. <https://doi.org/10.1007/s11528-021-00637-1>
25. Chiu, T. K. F. ve Chai, C.-s. (2020). Sustainable curriculum planning for artificial intelligence education: a self-determination theory perspective. *Sustainability*, 12(14), 5568. <https://doi.org/10.3390/su12145568>
26. Chiu, T. K., Xia, Q., Zhou, X., Chai, C. S. ve Cheng, M. (2023). Systematic literature review on opportunities, challenges, and future research recommendations of artificial intelligence in education. *Computers and Education: Artificial Intelligence*, 4, 100118. <https://doi.org/10.1016/j.caeai.2022.100118>.
27. *Christopoulos, A., Kajasilta, H., Salakoski, T. ve Laakso, M. J. (2020). Limits and virtues of educational technology in elementary school mathematics. *Journal of Educational Technology Systems*, 49(1), 59-81. <https://doi.org/10.1177/0047239520908838>
28. *Chu, H. C., Chen, J. M., Kuo, F. R. ve Yang, S. M. (2021). Development of an adaptive game-based diagnostic and remedial learning system based on the concept-effect model for improving learning achievements in mathematics. *Educational Technology & Society*, 24(4), 36-53. https://www.j-ets.net/collection/published-issues/24_4

29. *Chu, Y.-S., Yang, H.-C., Tseng, S.-S. ve Yang, C.-C. (2014). Implementation of a model-tracing-based learning diagnosis system to promote elementary students' learning in mathematics. *Educational Technology & Society*, 17(2), 347–357. <https://www.jstor.org/stable/jeductechsoci.17.2.347>
30. Cohen, J. A. (1960). Coefficient of agreement for nominal scales. *Education and Psychological Measurement*, 20, 37-46.
31. Cooper, H., Hedges, L. V. ve Valentine, J. C. (2009). *The handbook of research synthesis and meta-analysis* (2. baskı.). Russell Sage Publication.
32. Creswell JW, Plano Clark VL, Gutmann M, Hanson W. (2003). Advanced mixed methods research designs. Tashakkori A, Teddlie C (Eds.) *Handbook of Mixed Methods in Social and Behavioral Research* içinde. (s. 209–240). Sage.
33. Creswell, J. W. ve Sözbilir, M. (2017). *Karma yöntem araştırmalarına giriş*. Pegem Akademi.
34. Dağyar, M. (2014). *Probleme dayalı öğrenmenin akademik başarıya etkisi: bir meta-analiz çalışması* [Yayımlanmamış yüksek lisans tezi]. Hacettepe Üniversitesi.
35. Doğan, O. ve Baloglu, N. (2020). Üniversite öğrencilerinin endüstri 4.0 kavramsal farkındalık düzeyleri. *TÜBAV Bilim Dergisi*, 13(1), 126-142.
36. Duval, S. ve Tweedie, R. (2000). A nonparametric “trim and fill” method of accounting for publication bias in meta-analysis. *Journal of the american statistical association*, 95(449), 89-98.
37. Ellis, R. D. ve Allaire, J. C. (1999). Modeling computer interest in older adults: The role of age, education, computer knowledge, and computer anxiety. *Human factors*, 41(3), 345-355.
38. Fitton, V. A., Ahmedani, B. K., Harold, R. D. ve Shifflet, E. D. (2013). The role of technology on young adolescent development: Implications for policy, research and practice. *Child and Adolescent Social Work Journal*, 30, 399-413. DOI 10.1007/s10560-013-0296-2
39. *Francis, K., Rothsuh, S., Poscente, D. ve Davis, B. (2022). Malleability of spatial reasoning with short-term and long-term robotics interventions. *Technology, Knowledge and Learning*, 27(3), 927-956. <https://doi.org/10.1007/s10758-021-09520-7>
40. Fu, S., Gu, H. ve Yang, B. (2020). The affordances of AI-enabled automatic scoring applications on learners' continuous learning intention: An empirical study in China. *British Journal of Educational Technology*, 51(5), 1674-1692. <https://doi.org/10.1111/bjet.12995>
41. Garrison, D. R. (2017). *E-learning in the 21st century: A community of inquiry framework for research and practice*. Routledge.
42. Ginsberg, M. (2012). *Essentials of artificial intelligence*. Morgan Kaufmann Publishers
43. Glass, G. V. (1976). Primary secondary and meta-analysis of research. *Educational Researcher*, 5, 3-8.
44. Goldie, J. G. S. (2016). Connectivism: A knowledge learning theory for the digital age? *Medical Teacher*, 38(10), 1064–1069. <https://doi.org/10.3109/0142159X.2016.1173661>
45. *González-Calero, J. A., Cózar, R., Villena, R., & Merino, J. M. (2019). The development of mental rotation abilities through robotics-based instruction: An experience mediated by gender. *British Journal of Educational Technology*, 50(6), 3198-3213. <https://doi.org/10.1111/bjet.12726>
46. Guerra, F.C.H. (2022). *A model for putting connectivism into practice in a classroom environment*. [Unpublished Master Thesis, Universidade Nova de Lisboa], NOVA Information Management School.
47. Güler, N. ve Gelbal, S. (2010). A study based on classic test theory and many facet Rasch model. *Eurasian Journal of Educational Research*, 38, 108-125.
48. Güneş, Z. (2021). *Fen bilimlerinde FETEMM (fen, teknoloji, matematik, mühendislik) uygulamalarının çoklu bütüncül yaklaşımıyla karşılaştırmalı analizi*[Yayımlanmamış yüksek lisan tezi].Kilis 7 Aralık Üniversitesi.
49. Hambleton, R. K. ve Swaminathan, H. (1985). Estimation of Item and Ability Parameters. *Item Response Theory: Principles and Applications* içinde (s. 125-150). Springer Netherlands.
50. Hamet, P. ve Tremblay, J. (2017). Artificial intelligence in medicine. *Metabolism Clinical and Experimental*, 69,36-40. DOI: 10.1016/j.metabol.2017.01.011
51. Higgins, J. P., Thompson, S. G., Deeks, J. J. ve Altman, D. G. (2003). Measuring inconsistency in meta-analyses. *Bmj*, 327(7414), 557-560.
52. Holmes, W., Bialik, M. ve Fadel, C. (2019). Artificial intelligence in Education. *Center for Curriculum Redesign*. doi: 10.1007/978-3-319-60013-0_107-1

53. Hoorn, J. F., Huang, I. S., Konijn, E. A. ve van Buuren, L. (2021). Robot Tutoring of Multiplication: Over One-Third Learning Gain for Most, Learning Loss for Some. *Robotics*, 10(1), 16. <https://doi.org/10.3390/robotics10010016>
54. Humble, N. ve Mozeliuss, P. (2019). Artificial Intelligence in Education-a Promise, a Threat or a Hype?. In European Conference on the Impact of Artificial Intelligence and Robotics 2019 (ECIAIR 2019), Oxford, UK 149– 156.
55. İlhan, M. (2016). Açık uçlu sorularla yapılan ölçmelerde klasik test kuramı ve çok yüzeyli Rasch modeline göre hesaplanan yetenek kestirimlerinin karşılaştırılması. *Hacettepe Üniversitesi Eğitim Fakültesi Dergisi*, 31(2), 346-368. DOI:10.16986/HUJE.2016015182.
56. Jain, P., Coogan, S. C., Subramanian, S. G., Crowley, M., Taylor, S. ve Flannigan, M. D. (2020). A review of machine learning applications in wildfire science and management. *Environmental Reviews*, 28(4), 478–505. <https://doi.org/10.1139/er-2020-0019>
57. Jia, J., He, Y. ve Le, H. (2020). *A multimodal human-computer interaction system and its application in smart learning environments*, in International Conference on Blended Learning, (Cham: Springer), 3–14. doi: 10.1007/978-3-030-51968-1_1
58. Kablan, Z., Topan, B., ve Erkan, B. (2013). Sınıf içi öğretimde materyal kullanımının etkililik düzeyi: bir meta-analiz çalışması. *Kuram ve Uygulamada Eğitim Bilimleri*, 13(3), 1629-1644.
59. **Kajiwara, Y., Matsuoka, A., Shinbo, F. (2023). Machine learning role playing game: Instructional design of AI education for age-appropriate in K-12 and beyond. *Computers and Education: Artificial Intelligence*, 5. <https://doi.org/10.1016/j.caeai.2023.100162>
60. Kaplan-Rakowski, R., Grotewold, K., Hartwick, P. ve Papin, K. (2023). Generative AI ve teachers' perspectives on its implementation in education. *Journal of Interactive Learning Research*, 34(2), 313-338. Retrieved January 13, 2025 from <https://www.learntechlib.org/primary/p/222363/>.
61. Ke, F. (2010). Examining online teaching, cognitive, and social presence for adult students. *Computers & Education*, 55(2), 808-820. <https://doi.org/10.1016/j.compedu.2010.03.013>
62. Kim NJ ve Kim MK (2022) Teacher's Perceptions of Using an Artificial Intelligence-Based Educational Tool for Scientific Writing, *Frontiers in Education*. 7:755914. doi: doi.org/10.3389/feduc.2022.755914
63. Koçel, T. (2017). Yönetim ve organizasyonda metodoloji ve güncel kavramlar. *İstanbul Üniversitesi İşletme Fakültesi Dergisi*, 46, 3-8.
64. Komalavalli, K., Hemalatha, R. ve Dhanalakshmi, S. (2020). A survey of artificial intelligence in smart phones and its applications among the students of higher education in and around Chennai City. *Shanlax International Journal of Education*, 8(3), 89-95. doi:10.34293/education.v8i3.2379
65. Köse, İ. A., Usta, H. G. ve Yandı, A. (2016). Sunum yapma becerilerinin çok yüzeyli rasch analizi ile değerlendirilmesi. *Abant İzzet Baysal Üniversitesi Eğitim Fakültesi Dergisi*, 16(4), 1853-1864.
66. Lawshe, C. H. (1975). A quantitative approach to content validity. *Personnel Psychology*, 28, 563–575.
67. Li, B., Hou, B., Yu, W., Lu, X. ve Yang, C. (2017). Applications of artificial intelligence in intelligent manufacturing: A review. *Frontiers of Information Technology & Electronic Engineering*, 18(1), 86-96. <https://doi.org/10.1631/FITEE.1601885>
68. Li, M. ve Su, Y. (2020). Evaluation of online teaching quality of basic education based on artificial intelligence. *International Journal of Emerging Technologies in Learning (ijET)*, 15(16), 147–161. <https://doi.org/10.3991/ijet.v15i16.15937>
69. *Li, X., Xiang, H., Zhou, X. ve Jing, H. (2023, June). An empirical study on designing STEM+ AI teaching to cultivate primary school students' computational thinking perspective. In *Proceedings of the 2023 8th International Conference on Distance Education and Learning* (pp. 1-7).
70. Liang, Y. ve Chen, L. (2018). Analysis of current situation, typical characteristics and development trend of artificial intelligence education application. *China Electrofication Education*, 2018(3), 24-30
71. *Lin, X. F., Wang, Z., Zhou, W., Luo, G., Hwang, G. J., Zhou, Y. ve Liang, Z. M. (2023). Technological support to foster students' artificial intelligence ethics: An augmented reality-based contextualized dilemma discussion approach. *Computers & Education*, 201. <https://doi.org/10.1016/j.compedu.2023.104813>
72. Linacre, J. M. (1993). *Generalizability theory and many facet Rasch measurement*. Annual Meeting of the American Educational Research Association <https://files.eric.ed.gov/fulltext/ED364573.pdf>

73. Linacre, J. M. (2008). A user's guide to winsteps, ministep Rasch-model computer programs, program manuel, 3.66, P.O. Box. 811322, Chicago IL 60681-1322.
74. Litan, A. (2021). Hype Cycle for Blockchain 2021; More Action than Hype. *Gartner*. <https://blogs.gartner.com/avivahlitan/2021/07/14/hype-cycle-for-blockchain-2021-moreaction-than-hype>
75. Luckin, R., Holmes, W., Griffiths, M. ve Forcier, L.B. (2016). *Intelligence Unleashed - an Argument for AI in Education*. CL Knowledge Lab.
76. Luo, D. (2018). Guide teaching system based on artificial intelligence. *International Journal of Emerging Technologies in Learning (ijET)*, 13(08), 90–102. <https://doi.org/10.3991/ijet.v13i08.9058>
77. Lynch, B. K. ve McNamara, T. F. (1998). Using G-theory and many-facet Rasch measurement in the development of performance assessments of the ESL speaking skills of immigrants. *Language Testing*, 15(2), 158- 180. DOI: 10.1177/026553229801500202.
78. Lynch, B. K. ve McNamara, T. F. (1998). Using G-theory and many-facet Rasch measurement in the development of performance assessments of the ESL speaking skills of immigrants. *Language Testing*, 15(2), 158-180.
79. Maqbool, F., Ansari, S. ve Otero, P., 2021. The role of artificial intelligence and smart classrooms during covid-19 pandemic and its impact on education. *Journal of Independent Studies and Research Computing*, 19(1), 7-14. <https://doi.org/10.31645/JISRC.41.19.1.2>.
80. Mayring, P. (2000). Qualitative content analysis. *A Companion to Qualitative Research*, 1(2), 159-176
81. MEB (2024). *Eğitimde Kullanılan Yapay Zeka Araçları: Öğretmen El Kitabı*. Çevrimiçi Kitap
82. Mijwil, M.M., Aggarwal, K., Mutar, D.S., Mansour, N. ve Singh, R.S.S., 2022. the position of artificial intelligence in the future of education: an overview. *Asian Journal of Applied Sciences*, 10(2), 102-108. <https://doi.org/10.24203/ajas.v10i2.6956>.
83. Miles, M, B. ve Huberman, A. M. (1994). *Qualitative data analysis: An expanded Sourcebook*. (2nd ed). Sage.
84. Moher D, Liberati A, Tetzlaff J, Altman DG; PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. 21;6(7). doi: 10.1371/journal.pmed.1000097.
85. Molina-Azorín, J. F., López-Gamero, M. D., Pereira-Moliner, J. ve Pertusa-Ortega, E. M. (2012). Mixed methods studies in entrepreneurship research: Applications and contributions. *Entrepreneurship & Regional Development*, 24(5-6), 425-456. <https://doi.org/10.1080/08985626.2011.603363>
86. Mu, P. (2019). *Research on artificial intelligence education and its value orientation*. Paper presented at the 1stInternational Education Technology and Research Conference (IETRC 2019, China, Retrieved from https://webofproceedings.org/proceedings_series/ESSP/IETRC%202019/IETRC19165.pdf
87. Nichols, J. A., Herbert Chan, H. W. ve Baker, M. A. (2019). Machine learning: Applications of artificial intelligence to imaging and diagnosis. *Biophysical reviews*, 11(1), 111–118. DOI: 10.1007/s12551-018-0449-9
88. Ouyang, F. ve Jiao, P. (2021). Artificial intelligence in education: The three paradigms. *Computers and Education: Artificial Intelligence*, 2, 100020.
89. Panigrahi, C. M. A. (2020). Use of artificial intelligence in education. *Management Accountant*, 55, 64-67. <https://ssrn.com/abstract=3606936>
90. *Pareto, L. (2014). A teachable agent game engaging primary school children to learn arithmetic concepts and reasoning. *International Journal of artificial intelligence in education*, 24, 251-283. DOI 10.1007/s40593-014-0018-8
91. Pedró, F., Subosa, M., Rivas, A. ve Valverde, P. (2019). Artificial intelligence in education: Challenges and opportunities for sustainable development. *UNESCO*
92. Qin, F., Li, K. ve Yan, J. (2020). Understanding user trust in artificial intelligence-based educational systems: evidence from China. *Br. J. Educ. Technol.* 51, 1693–1710. doi: 10.1111/bjet.12994
93. Randall, N. (2019). A Survey of Robot-Assisted Language Learning (RALL). *ACM Transactions on Human-Robot Interaction*, 9(1), 1-36. <https://doi.org/10.1145/3345506>
94. *Rau, M. A., Aleven, V. ve Rummel, N. (2017). Making connections among multiple graphical representations of fractions: Sense-making competencies enhance perceptual fluency, but not vice versa. *Instructional Science*, 45, 331-357. DOI 10.1007/s11251-017-9403-7

95. *Rau, M. A., Aleven, V., Rummel, N. ve Pardos, Z. (2014). How should intelligent tutoring systems sequence multiple graphical representations of fractions? A multi-methods study. *International Journal of Artificial Intelligence in Education*, 24, 125-161. DOI 10.1007/s40593-013-0011-7
96. Reeves, B., Hancock, J. ve Liu, X. (2020). Social Robots Are Like Real People: First Impressions, Attributes, and Stereotyping of Social Robots. *Technology, Mind, and Behavior*, 1(1). <https://doi.org/10.1037/tmb0000018>
97. Rodríguez, M.D.(2001).Glossary on meta-analysis. *J.Epidemiol Community Health*, 55(8). doi: 10.1136/jech.55.8.534.
98. Roscoe, R. D., Walker, E. A. ve Patchan, M. M. (2018). Facilitating peer tutoring and assessment in intelligent learning systems. In S. D. Craig (Ed.), *Tutoring and intelligent tutoring systems* (pp. 41-68). Nova Science Publishers, Inc.
99. Rosenthal, R. (1979). The file drawer problem and tolerance for null results. *Psychological Bulletin*, 86(3), 638–641. <https://doi.org/10.1037/0033-2909.86.3.638>
100. Russell, S. J. ve Norvig, P. (2009). *Artificial Intelligence: A Modern Approach* (3rd ed.). Prentice Hall.
101. Ryu, M. ve Han, S. (2018). The educational perception on artificial intelligence by elementary school teachers. *J. Korean Assoc. Inform. Educ.* 22, 317–324. doi: 10.1016/j.ridd.2019.01.002
102. Sak, R., Sak, İ. T. Ş., Şendil, Ç. Ö. Ve Nas, E. (2021). Bir araştırma yöntemi olarak doküman analizi. *Kocaeli Üniversitesi Eğitim Dergisi*, 4(1), 227-256. <https://doi.org/10.33400/kuje.843306>
103. **Salas-Pilco, S.Z. (2020), The impact of AI and robotics on physical, social-emotional and intellectual learning outcomes: An integrated analytical framework. *Br J Educ Technol*, 51: 1808-1825. <https://doi.org/10.1111/bjet.12984>
104. Santos, O. C. (2016). Training the body: The potential of AIED to support personalized motor skills learning. *International Journal of Artificial Intelligence in Education*, 26(2), 730-755. <https://doi.org/10.1007/s40593-016-0103-2>
105. Schmidt, F. L., Oh, I. S. ve Hayes, T. L. (2009). Fixed-versus random-effects models in meta-analysis: Model properties and an empirical comparison of differences in results. *British journal of mathematical and statistical psychology*, 62(1), 97-128. doi: 10.1348/000711007X255327.
106. Self, J. (2016). The birth of IJAIED. *International Journal of Artificial Intelligence in Education*, 26(1), 4-12. <https://doi.org/10.1007/s40593-015-0040-5>
107. Semerci, Ç. (2011). Mikro öğretim uygulamalarının çok-yüzeyle Rasch ölçme modeli ile analizi. *Eğitim ve Bilim*, 36(161). <https://egitimvebilim.ted.org.tr/index.php/EB/article/view/145> adresinden erişildi.
108. Semerci, Ç. (2012). Öğrencilerin BÖTE bölümüne ilişkin görüşlerinin rasch ölçme modeline göre değerlendirilmesi (Fırat Üniversitesi örneği). *Education Sciences*, 7(2), 777-784.
109. Shamir, G. ve Levin, I. (2021).Teaching machine learning in elementary school. *International Journal of Child-Computer Interaction*.31 (2022). <https://doi.org/10.1016/j.ijcci.2021.100415>
110. **Shamir, G. ve Levin, I. (2024). *Exploring a practical implementation of teaching machine learning in elementary schools*.<http://dx.doi.org/10.2139/ssrn.4752354>
111. Sharma, R., Kamble, S. S., Gunasekaran, A., Kumar, V. ve Kumar, A. (2020). A systematic literature review on machine learning applications for sustainable agriculture supply chain performance . *Computers and Operations Research*, 119. <https://doi.org/10.1016/j.cor.2020.104926>Haklar ve içerik edinin
112. Siemens, G. (2005). Connectivism: A learning theory for the digital age. *International Journal of Instructional Technology and Distance Learning*, 2. http://www.itdl.org/Journal/Jan_05/article01.htm
113. Sterne, J. A. ve Harbord, R. M. (2004). Funnel plots in meta-analysis. *The Stata Journal: Promoting Communications on Statistics and Stata*, 4(2), 127–141. <https://doi.org/10.1177/1536867X0400400204>
114. Streubert, H. J. ve Carpenter, D. R. (2011). *Qualitative research in nursing: Advancing the humanistic imperative*. Wolters Kluwer.
115. Su, J. ve Yang, W. (2022). Artificial intelligence in early childhood education: A scoping review. *In Computers and education: Artificial Intelligence*, 3. <https://doi.org/10.1016/j.caeai.2022.100049>
116. Talan, T. (2021). Artificial intelligence in education: A bibliometric study. *International Journal of Research in Education and Science (IJRES)*, 7(3), 822-837. <https://doi.org/10.46328/ijres.2409>
117. Talan, H. (2020). *Jigsaw tekniğinin çoklu bütüncül yaklaşımla analizi bağlamında fen bilgisi öğretim programında kullanılmasının değerlendirilmesi* [Yayımlanmamış yüksek lisans tezi]. Kilis 7 Aralık Üniversitesi.

118. Thalheimer, W. ve Cook, S. (2002). How to calculate effect sizes from published research: A simplified methodology. *Work-Learning Research*, 1(9), 1-9.
119. Toivonen, T., Jormanainen, I., Kahila, J., Tedre, M., Valtonen, T. ve Vartiainen, H. (2020, July). Co-designing machine learning apps in K-12 with primary school children. In *2020 IEEE 20th International Conference on Advanced Learning Technologies (ICALT)* (pp. 308-310). IEEE.
120. Topol, E. (2019). Deep medicine: How artificial intelligence can make healthcare human again. Basic Books.
121. Toraman, S. (2021). Karma Yöntemler Araştırması: Kısa Tarihi, Tanımı, Bakış Açıları ve Temel Kavramlar. *Nitel Sosyal Bilimler*, 3(1), 1-29. <https://doi.org/10.47105/nsb.847688>
122. Tsagris, M. ve Fragkos, K. C. (2018). Meta-analyses of clinical trials versus diagnostic test accuracy studies. *Diagnostic meta-analysis: A useful tool for clinical decision-making*, 31-40.
123. Uluman, M. ve Tavşancıl, E. (2017). Çok değişkenlik kaynaklı rasch ölçme modeli ve hiyerarşik puanlayıcı modeli ile kestirilen puanlayıcı parametrelerinin karşılaştırılması. *İnsan ve Toplum Bilimleri Araştırmaları Dergisi*, 6(2), 777-798. <https://doi.org/10.15869/itobiad.296489>
124. UNESCO (2018). Öğretmenlere Yönelik Bilgi ve İletişim Teknolojileri Yetkinlik Çerçevesi. <https://yegitek.meb.gov.tr/www/unesco-ogretmenlere-yonelik-bilgi-ve-iletisim-teknolojileri-yetkinlik-cercevesi/icerik/3146>
125. UNESCO (2019). *Artificial Intelligence in Education: Challenges and Opportunities for Sustainable Development*. <https://unesdoc.unesco.org/ark:/48223/pf0000366994>
126. Veneziano L. ve Hooper J. (1997). A method for quantifying content validity of health-related questionnaires. *American Journal of Health Behavior*, 21(1), 67-70.
127. ***Wang, X., Liu, Q., Pang, H., Tan, S. C., Lei, J., Wallace, M. P. ve Li, L. (2023). What matters in AI-supported learning: A study of human-AI interactions in language learning using cluster analysis and epistemic network analysis. *Computers & Education*, 194. <https://doi.org/10.1016/j.compedu.2022.104703>
128. *Wang, X., Pang, H., Wallace, M. P., Wang, Q. ve Chen, W. (2024). Learners' perceived AI presences in AI-supported language learning: A study of AI as a humanized agent from community of inquiry. *Computer Assisted Language Learning*, 37(4), 814-840. DOI: 10.1080/09588221.2022.2056203
129. Wang, Y.-C. (2014). Using wikis to facilitate interaction and collaboration among EFL learners: A social constructivist approach to language. *System*, 42, 383-390. doi:10.1016/j.system.2014.01.007
130. Wang, S. P. ve Chen, Y. L. (2018). Effects of multimodal learning analytics with concept maps on college students' vocabulary and reading performance. *Journal of Educational Technology & Society*, 21(4), 12-25.
131. Yang, W. (2022). Artificial Intelligence education for young children: Why, what, and how in curriculum design and implementation. *Computers and Education: Artificial Intelligence*, 3, <https://doi.org/10.1016/j.caeai.2022.100061>
132. Yue, M., Jong, M. S. Y., & Dai, Y. (2022). Pedagogical design of K-12 artificial intelligence education: A systematic review. *Sustainability*, 14(23). <https://doi.org/10.3390/su142315620>
133. Yurdugül, H. (2005, Eylül). Ölçek Geliştirme Çalışmalarında Kapsam Geçerliği İçin Kapsam Geçerlik İndekslerinin Kullanılması. XIV. Ulusal Eğitim Bilimleri Kongresi, Pamukkale Üniversitesi Eğitim Fakültesi, Denizli.
134. Zhang, K. ve Aslan, A. B. (2021). AI technologies for education: Recent research & future directions. *Computers and Education*, 2. <https://doi.org/10.1016/j.caeai.2021.100025>.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.