

The Effects of Supplementary Artificial Intelligence Courseware Use on Elementary Students' Scientific Inquiry Skills and Scientific Communication Skills

Soyeon Kim*

Yeoul Elementary School, Sejong, Republic of Korea

*Corresponding Author: yeon.y083@gmail.com

ABSTRACT

This study examined the effects of supplementary learning using artificial intelligence (AI) courseware, implemented in conjunction with regular elementary science instruction, on the scientific inquiry skills and scientific communication skills of students. The participants were 111 sixth-grade students from six classes in an elementary school in S City, who were assigned to an experimental group (three classes, 57 students) and a comparison group (three classes, 54 students). Both groups completed the same 32-lesson science curriculum over one semester. In addition to regular instruction, the experimental group participated in supplementary learning based on individual AI courseware for 20 min, 3 times a week, while the comparison group participated in supplementary learning using textbooks, experiment observation materials, and problem-solving worksheets for the same duration. The pre-test and post-test data were analyzed using analysis of covariance, with the pre-test scores treated as covariates. The results did not reveal statistically significant group differences in overall scientific inquiry skills. However, significant group differences were observed in specific subdomains of scientific communication skills, particularly in the explanatory purposes of “What” and “How” and in the numerical form of communication. These findings suggest that the effects of the use of additional AI courseware are limited and domain-specific. Although AI courseware may support numerical information processing and foundational explanatory communication skills, it appears insufficient to foster higher-order inquiry skills and more advanced forms of scientific communication without deliberate instructional mediation.

KEY WORDS: Artificial intelligence courseware; elementary science education; scientific communication skills; scientific inquiry skills; supplementary learning

INTRODUCTION

Digital textbooks in Korea have been developed and distributed under the supervision of the Ministry of Education since the late 1990s, with large-scale implementation in schools beginning in 2014. Previous studies have reported that digital textbooks can improve students' academic achievement compared to traditional printed textbooks (Song and Park, 2009; Zucker et al., 2009) and have positive effects on problem-solving skills (Choi and Sim, 2024; Choi and Seo, 2009; Seo et al., 2009). However, despite these reported benefits, research has also identified several limitations associated with the use of digital textbooks. Findings on their effects on self-directed learning remain inconsistent, with both positive and negative results reported (Joo and Lim, 2015; Yoo et al., 2025; Choi and Sim, 2024). In addition, some studies have suggested that digital textbook-based instruction can constrain the development of higher-order thinking skills (Ahn and Kim, 2025). Moreover, increased exposure to digital textbooks does not necessarily lead to improved academic achievement (Choppin et al., 2014).

In response to these limitations, recent discussions have emphasized a shift from content-delivery-oriented digital

textbooks to systems that provide personalized support based on learners' characteristics and performance. The Korean Ministry of Education has promoted the development of Artificial Intelligence-based Digital Textbooks (AIDTs) with the goal of implementing personalized education, highlighting learner diagnosis and adaptive support as core functions (Lee and Lee, 2024). AIDTs are conceptualized as adaptive learning platforms that integrate textbook content, learning management, and assessment functions. In particular, the use of learner performance data to diagnose learning levels and provide tailored tasks and feedback closely resembles the functions of commercially available adaptive learning courseware (Kim and Kim, 2025; Lee and Lee, 2024). Although policy initiatives have focused on AIDTs, in actual school settings, these trends have often been realized earlier through the adoption of commercial Artificial intelligence (AI)-based courseware.

AI-based courseware has been increasingly introduced in school subjects, including science, through government-led initiatives and the operation of digital-leading schools. In elementary schools, in particular, commercial AI courseware has been widely adopted using budgets allocated for projects

related to digital textbooks. Courseware, a compound of “course” and “software,” refers to educational software designed to support curriculum content, learning processes, and instructional strategies. AI courseware uses artificial intelligence technologies to diagnose the current levels of learners, provide personalized learning content, and evaluate learning processes (Son and Kang, 2024). In this study, AI courseware is conceptualized as a type of intelligent tutoring system that structures curriculum content and assessment elements while adapting learning paths, difficulty levels, and feedback based on learners’ performance data (Shin and Hwang, 2024; Lee, 2025).

Internationally, AI-based courseware and intelligent tutoring systems have been discussed as key tools to implement personalization and adaptive learning in science education. Holmes et al. (2019) reported that AI-supported learning environments can contribute to the understanding of complex scientific concepts and the development of higher-order thinking skills by adjusting task difficulty and feedback based on learner data. Similarly, Means et al. (2014) found that well-designed technology-based instruction in online and digital learning environments can have positive effects on academic achievement compared to traditional instruction. Despite these discussions, empirical studies that examine the effects and limitations of AI courseware specifically within science education contexts remain relatively limited.

Within the Korean context, research on AI courseware has focused mainly on subjects such as Korean language, English, and mathematics. Studies in Korean language education have reported increased student interest and reduced learning burden through personalized content and support for foundational skills (Lee and Koo, 2024), as well as improvements in reading ability and reduced achievement gaps (Lim, 2024). In English education, positive effects have been observed on level-differentiated reading skills (Lee, 2024). In mathematics education, AI-based digital tools have been shown to support a shift from problem-solving-centered instruction toward inquiry- and expression-oriented learning by providing diverse digital learning experiences, including computational practice, game-based activities, and manipulative tasks (Hoyles, 2018; Bray and Tangney, 2017; Hoyles and Lagrange, 2009).

Nevertheless, the reported benefits of AI courseware and digital instructional resources have not been consistent across all learning domains. Choppin et al. (2014) noted limitations in the development of mathematical communication skills through digital curricular materials. Other studies have raised concerns about the mismatches between algorithm-driven interventions and learners’ characteristics (Choi et al., 2024), potential biases embedded in recommendation algorithms (Jang, 2024), and the risk that overreliance on AI-generated guidance can hinder learners’ creative and critical thinking (Lee and Gu, 2024).

Science education, by its nature, aims not only to transmit scientific knowledge but also to cultivate scientific thinking through inquiry and communication. The 2022 Revised

Science Curriculum in Korea explicitly identifies scientific inquiry skills and scientific communication skills as core competencies (Ministry of Education, 2022). Scientific communication involves interpreting and justifying results obtained through inquiry using scientific language, coordinating understanding through interaction with others, and collaboratively solving problems (Lemke, 1990). Approximately 11% of science curriculum achievement standards are directly related to communication (Jeon, 2019), underscoring its essential role in science learning. However, commercially available AI courseware commonly used in schools is often designed around individualized problem-solving and automated feedback, which may not fully align with the inquiry-oriented and communicative goals of science education.

The AI courseware examined in this study does not directly provide traditional inquiry activities such as hands-on experimentation or data collection. Instead, it includes tasks that require learners to apply scientific concepts, interpret situations, select or justify explanations, and analyze information presented in numerical, graphical, or symbolic forms. Through these tasks, students can engage in certain components of scientific inquiry and scientific communication during problem-solving processes. However, it remains unclear whether such item-based experiences translate into measurable improvements in inquiry or communication skills.

Therefore, this study aims to examine the effects of the use of additional AI courseware in an elementary science context on the scientific inquiry skills and scientific communication skills of students. Specifically, this study investigates whether AI courseware-based supplementary learning influences overall inquiry skills and distinct subdomains of scientific communication, thus providing empirical evidence on both the educational potential and limitations of AI courseware in elementary science education.

METHODOLOGY

Participants and Research Design

The participants in this study were 111 sixth-grade students from six classes at an elementary school located in S City. The students were assigned to an experimental group consisting of three classes (57 students) and a comparison group consisting of three classes (54 students). The assignment of the group followed the existing organization of the class, as the random assignment was not feasible in the school context. To control for teacher-related variables, the same teacher taught science lessons for both groups.

Both groups completed the same regular science curriculum over one semester, consisting of 32 lessons taught 3 times a week. The instructional units included Light and Lenses (10 lessons), Uses of Electricity (11 lessons), and Seasonal Changes (11 lessons), which were taught within the same semester according to curriculum scheduling. During regular class time, instruction was conducted using textbooks and

experiment-observation materials, and students engaged in inquiry and investigation activities aligned with each unit.

In addition to regular instruction, supplementary learning sessions were conducted 3 times a week during morning hours, each session lasting 20 min. The experimental group participated in individual supplementary learning using AI courseware aligned with the content covered in regular science lessons. During each session, students solved problems related to the unit most recently taught and received immediate feedback, including correctness and explanatory solutions. The AI courseware presented items that require students to explain observed phenomena, select or justify reasons and processes, and interpret or compare the data presented in tables, graphs, and diagrams. After completing a set of items, the system diagnosed the learning levels of the students and provided adaptive follow-up tasks or additional learning activities.

The comparison group engaged in supplementary learning for the same amount of time using science textbooks, experiment-observation materials, and teacher-provided worksheets. The students worked individually or in pairs to review the previously learned content. The teacher circulated during these sessions, reviewed the responses of the students, and provided oral feedback. Thus, regular science instruction was identical for both groups, and the primary difference between the groups was the use of AI courseware during supplementary learning.

Given the practical constraints of the school setting, a quasi-experimental pre-test–post-test design was used (Campbell and Stanley, 2015). The overall research design is summarized in Table 1.

Instruments

To examine the effects of AI courseware–based supplementary learning, quantitative data were collected using a pre-test–post-test design. Students’ scientific inquiry skills were measured using the Scientific Inquiry Skills Test developed by Kwon and Kim (1994). This instrument assesses both basic and integrated inquiry skills and originally consists of 15 multiple-choice items. Taking into account changes in students’ experiences and terminology since the instrument was developed, two items related to observation and measurement were excluded after expert review. The final version included 13 items.

The validity of the content was examined by a professor of science education, two researchers holding doctoral degrees in science education, and three elementary school teachers with over 10 years of teaching experience and master’s degrees in science education. Items were rated appropriate (1) or inappropriate (0), and the content validity index ranged from 0.83 to 0.92, indicating acceptable validity. Reliability analysis based on pre-test data yielded a Cronbach’s α of 0.74.

Scientific communication skills were assessed using the Scientific Communication Skills Test developed by Jeon (2019). In this study, scientific communication was defined as the ability to convey, exchange, and share scientific explanations and claims about facts, phenomena, and causes

using scientific knowledge and literacy. The instrument consists of 24 items, including 16 multiple-choice items and 8 constructed-response items, designed for students in Grades 3–6.

Items are categorized by communicative purpose – scientific explanation (What, How, Why) and scientific argumentation (Claim, Evidence, Rebuttal) – and by communication form (text, number, table, and diagram), with six items in each form category. The definitions of each communication purpose are presented in Table 2. Constructed-response items were scored using a three-point rubric (2 = correct, 1 = partially correct, 0 = incorrect), with predefined scoring criteria. Six experts independently scored the responses, and discrepancies were resolved through discussion to ensure reliability of the scoring. Reliability analysis of the objective items based on the pre-test data showed high internal consistency (Cronbach’s α = 0.88).

Characteristics of the AI Courseware Used in the Study

In this study, the AI courseware used for the experimental treatment was selected through a systematic review process based on established analytical frameworks. Specifically, the analytical framework for digital curricular resources proposed by Choppin et al. (2014) and the AI courseware typology developed by Son and Kang (2024) were used as foundational references. Based on these studies, the analytical framework was reconstructed to focus on three key dimensions relevant

Table 1: Research design

Group	Pretest	Treatment	Posttest
Experimental group	O1	X1	O2
Comparison group	O3	X2	O4

O1 and O3 indicate the pre-tests of scientific inquiry skills and scientific communication skills. X1 represents supplementary learning using AI courseware. X2 represents additional learning using textbooks, experiment–observation materials, and worksheets. O2 and O4 indicate the post-tests of scientific inquiry skills and scientific communication skills

Table 2: Definitions of domains according to the purposes of scientific communication

Domain (maximum score)	Description
Explanation	
What (3)	Describing what occurs in a phenomenon by identifying its features, data, or results without establishing causal relationships.
How (3)	Describe how a phenomenon occurs by describing processes, sequences, or mechanisms, partially including causal relationships.
Why (3)	Explain why a phenomenon occurs by providing causal reasoning and scientific principles.
Argumentation	
Claim (3)	Developing a claim or conclusion based on one’s own reasoning or interpretation.
Evidence (3)	Supporting a claim using scientific evidence or data.
Rebuttal (3)	Responding to or challenging a claim by presenting counterarguments or alternative explanations.

to elementary science instruction: learning experience, individualized learning, and communication.

Using this reconstructed framework, three commercially available AI courseware programs applicable to elementary science education were identified as preliminary candidates and labeled C1, C2, and C3. The functionality of each courseware program was examined according to the detailed criteria of the framework, and a comparative analysis of their features is presented in Table 3.

Among the candidates, C1 provided comprehensive support in all three dimensions. It supported learning experiences through video materials, problem presentation, and immediate feedback with explanatory solutions. In terms of personalized learning, C1 diagnosed the level of learning of the students based on performance data and automatically suggested the appropriate follow-up tasks and learning activities. In addition, C1 included multiple communication features, such as messaging, email, discussion boards, and chatbot functions, which facilitated interaction between teachers and students.

In contrast, C2 was designed to operate on proprietary devices, which limited its applicability in typical school conditions. Taking into account the need for feasibility in real classroom settings and balanced support for learning experience, individualized learning, and communication – while allowing flexible integration with regular classroom instruction and teacher mediation – C1 was selected as AI courseware for the experimental intervention in this study.

Ethical Considerations

Participation in this study was voluntary, and informed consent was obtained from all participants before data collection. The study was conducted as part of regular classroom instruction and involved non-interventional educational activities using anonymized data.

Data Collection and Analysis

A pre-test–post-test design was used to collect data. The group (experimental vs. comparison) served as an independent variable, and the post-test scores were treated as the dependent variable. Pre-test scores were included as covariates. One-way analysis of covariance (ANCOVA) was

conducted to examine group differences while controlling for initial differences.

Before analysis, ANCOVA assumptions were examined, including homogeneity of variances using the Levene test and the interaction between covariates and independent variables. Data were analyzed using IBM Statistical Package for the Social Sciences Statistics 27.0. Effect sizes were reported using partial eta squared (partial η^2), with values of 0.01, 0.06, and 0.14 interpreted as small, medium, and large effects, respectively (Cohen, 2013).

RESULTS

Scientific Inquiry Skills

To examine the effects of AI courseware–based supplementary learning on students' scientific inquiry skills, ANCOVA was conducted using post-test scores as the dependent variable and pre-test scores as covariates. Preliminary analyses indicated that there were no statistically significant differences between the experimental and comparison groups on the pre-test scores for scientific inquiry skills ($p > 0.05$), confirming baseline equivalence between the groups. The assumptions for ANCOVA, including homogeneity of variances and normality, were satisfied.

The ANCOVA results showed that the group did not have a significant main effect on basic inquiry skills, $F(1, 107) = 1.21, p = 0.275$, or integrated inquiry skills, $F(1, 107) = 0.00, p = 0.989$ (Tables 4 and 5). The effect sizes (partial η^2) for both outcomes were negligible, indicating that supplementary learning using AI courseware did not produce statistically significant improvements in general scientific inquiry skills compared to traditional supplementary learning.

Further analyses were conducted to examine group differences between subdomains of scientific inquiry skills. As shown in Tables 6 and 7, no statistically significant group effects were observed for any of the basic inquiry skill subdomains

Table 3: Comparative analysis of AI courseware features

Dimension	Feature	C1	C2	C3
Learning experience	Text and images		○	
	Video materials	○	○	○
	Presentation of problems	○	○	○
	Interactive activities		○	○
Individualized learning	Program-driven learning enforcement			○
	Program-driven learning recommendations	○	○	
	Teacher-driven learning recommendations		○	○
Communication	Messaging and email	○	○	○
	Comment writing			
	Discussion board	○		○
	Chatbot	○		○

Table 4: ANCOVA results for basic inquiry skills

Source	Sum of squares	df	Mean square	F	p
Pre-test (Covariate)	12.16	1	12.16	2.27	0.135
Group	6.45	1	6.45	1.21	0.275
Error	572.27	107	5.35		
Total	4304.00	110			

Table 5: ANCOVA results for integrated inquiry skills

Source	Sum of squares	df	Mean square	F	p
Pre-test (Covariate)	5.80	1	5.80	1.41	0.239
Group	0.00	1	0.00	0.00	0.989
Error	451.67	107	4.13		
Total	457.00	110			

(observation, classification, measurement, inference, and prediction) or the integrated inquiry skill subdomains (data transformation, data interpretation, hypothesis setting, variable control, and generalization). All partial η^2 values were small or close to zero, suggesting a minimal practical impact of the intervention on specific components of inquiry skills.

In general, these results indicate that AI courseware-based supplementary learning did not lead to significant gains in students' scientific inquiry skills at the general or subdomain levels.

Scientific Communication Skills

ANCOVA was also conducted to examine the effects of supplementary learning based on AI courseware on scientific communication skills. Pre-test scores were included as covariates to control for initial differences. The analysis did not reveal a significant main effect of the group on the total scientific communication scores, $F(1, 107) = 0.21, p = 0.650$. However, the covariate (pre-test score) showed a significant effect, $F(1, 107) = 1184.11, p < 0.001$, indicating a strong relationship between the performance of the pre-test and the post-test (Table 8).

To further explore potential differential effects, ANCOVA was conducted for subdomains based on communicative

purpose. As presented in Table 9, statistically significant group effects were found for the explanatory subdomains "What" ($p = 0.002$) and "How" ($p = 0.017$), the experimental group showing adjusted mean scores higher than those of the comparison group. The effect sizes for these subdomains ranged from small to modest, suggesting that AI courseware-based supplementary learning had a positive effect on students' ability to select and structure explanatory content at a foundational level. In contrast, no significant group effects were observed for the explanatory subdomain "Why" or any of the argumentation subdomains (Claim, Evidence, Rebuttal), and the corresponding effect sizes were minimal.

Additional analyses by communication form revealed a significant group effect only in the "Number" subdomain, $F(1, 107) = 10.25, p = 0.002$, with a medium effect size (partial $\eta^2 = 0.084$), as shown in Table 10. The experimental group demonstrated higher adjusted mean scores than the comparison group in numerical expression. No significant group differences were found in the text, table, or diagram subdomains ($p > 0.05$), and the effect sizes for these forms were small or negligible.

Taken together, these results indicate improvements in limited aspects of scientific communication, particularly in numerical expression and basic explanatory purposes, whereas

Table 6: ANCOVA results by domains of basic inquiry skills

Domain	Group	Adjusted mean	SD	F	p	partial η^2
Observation	Experimental (n=57)	1.37	0.56	1.38	0.240	<0.001
	Comparison (n=54)	1.37	0.49			
Classification	Experimental (n=57)	1.63	0.72	1.72	0.301	0.010
	Comparison (n=54)	1.70	0.75			
Measurement	Experimental (n=57)	1.12	0.73	1.13	0.126	0.019
	Comparison (n=54)	1.06	0.89			
Inference	Experimental (n=57)	1.60	0.82	1.57	0.210	0.010
	Comparison (n=54)	1.56	0.84			
Prediction	Experimental (n=57)	1.86	0.61	1.85	0.170	<0.001
	Comparison (n=54)	1.78	0.66			

ANCOVA: Analysis of covariance

Table 7: ANCOVA results by domains of integrated inquiry skills

Domain	Group	Adjusted mean	SD	F	p	partial η^2
Data transformation	Experimental (n=57)	1.26	0.58	1.26	0.729	<0.001
	Comparison (n=54)	1.28	0.60			
Data interpretation	Experimental (n=57)	1.19	0.86	1.16	0.617	0.002
	Comparison (n=54)	1.19	0.78			
Hypothesis setting	Experimental (n=57)	1.19	0.86	1.22	0.467	0.002
	Comparison (n=54)	1.19	0.78			
Variable control	Experimental (n=57)	1.44	0.85	1.46	0.580	0.003
	Comparison (n=54)	1.41	0.79			
Data generalization	Experimental (n=57)	1.11	0.92	1.11	0.290	<0.001
	Comparison (n=54)	1.09	0.90			
Prediction	Experimental (n=57)	1.42	0.73	1.42	0.633	0.002
	Comparison (n=54)	1.35	0.78			

ANCOVA: Analysis of covariance

Table 8: ANCOVA results for scientific communication ability

Source	Sum of squares	df	Mean square	F	p
Pre-test (Covariate)	3723.08	1	3723.08	1184.11	<0.001*
Group	0.65	1	0.65	0.21	0.650
Error	336.43	107			
Total	6613.09	110			

*p<0.05. ANCOVA: Analysis of covariance

Table 9: ANCOVA results by domains of communication purpose

Domain (Purpose)	Group	Adjusted mean	SD	F	p	partial η^2
Explanation						
What	Experimental (n=57)	4.83	1.99	4.83	0.002*	0.046
	Comparison (n=54)	4.00	1.90			
How	Experimental (n=57)	3.58	1.77	5.87	0.017*	0.052
	Comparison (n=54)	2.91	1.72			
Why	Experimental (n=57)	3.24	1.61	0.32	0.570	0.002
	Comparison (n=54)	2.93	1.39			
Argumentation						
Claim	Experimental (n=57)	4.71	1.92	2.78	0.096	0.035
	Comparison (n=54)	4.78	2.01			
Evidence	Experimental (n=57)	5.53	2.47	0.28	0.597	0.003
	Comparison (n=54)	5.42	2.05			
Rebuttal	Experimental (n=57)	3.67	2.42	0.18	0.673	0.002
	Comparison (n=54)	3.78	2.45			

*p<0.05. ANCOVA: Analysis of covariance, SD: Standard deviation

Table 10: ANCOVA results by domains of communication form

Communication form	Group	Adjusted mean	SD	F	p	partial η^2
Text	Experimental (n=57)	6.74	2.56	0.32	0.593	0.003
	Comparison (n=54)	6.59	2.51			
Number	Experimental (n=57)	5.95	3.15	10.25	0.002*	0.084
	Comparison (n=54)	5.52	2.69			
Table	Experimental (n=57)	5.79	2.06	0.32	0.437	0.007
	Comparison (n=54)	5.31	2.73			
Diagram	Experimental (n=57)	6.66	3.18	0.03	0.857	<0.001
	Comparison (n=54)	6.64	3.10			

*p<0.05. ANCOVA: Analysis of covariance, SD: Standard deviation

little change was observed in higher-order explanatory or argumentative communication skills.

DISCUSSION

The absence of significant effects on students' scientific inquiry skills may be explained by a mismatch between the nature of inquiry skills and the way AI courseware was used in this study. Scientific inquiry skills involve generating investigable questions, designing procedures, controlling variables, and interpreting evidence through open-ended and iterative processes. However, in the present intervention, the AI courseware primarily provided item-based problem-solving tasks with predefined answer spaces and automated feedback. Although such tasks may reinforce conceptual understanding and support procedural practice, they offer

limited opportunities for students to engage in the full cycle of inquiry, including hypothesis generation and experimental design. This misalignment between the targeted competencies and the instructional affordances emphasized during implementation may explain the negligible effects observed in both basic and integrated inquiry skills.

The limited effects observed in higher-order explanatory communication (Why) and scientific argumentation (Claim, Evidence, Rebuttal) may be attributed to the interactional characteristics of the AI courseware used in this study. Scientific argumentation and causal explanation are inherently dialogic and social processes that require opportunities for students to compare ideas, justify claims, respond to counterarguments, and negotiate meaning through interaction. However, AI courseware primarily emphasized personalized

learning with automated feedback focused on correctness, providing limited opportunities for epistemic dialog or peer-to-peer reasoning. As a result, while the courseware supported basic explanatory communication, it was less effective in fostering communicative practices that depend on sustained interaction and argumentative exchange.

CONCLUSION AND IMPLICATIONS

This study examined the effects of AI courseware-based supplementary learning on elementary students' scientific inquiry skills and scientific communication skills. A total of 111 sixth-grade students (57 in the experimental group and 54 in the comparison group) participated in the study, and the intervention was implemented during the supplementary learning phase following regular science instruction. The results indicated that additional learning using AI courseware did not lead to statistically significant improvements in the overall scientific inquiry skills of students. With respect to scientific communication skills, significant improvements were observed only in limited subdomains, particularly within specific forms and purposes of communication.

These findings suggest that the use of AI courseware alone in supplementary learning contexts has limitations in fostering a comprehensive development of scientific inquiry skills and scientific communication skills. Although AI courseware may support certain foundational aspects of communication, its impact appears to be constrained when it is primarily used for problem-solving and concept reinforcement without sufficient opportunities for extended inquiry and interaction.

Based on these results, several implications can be drawn for the effective use of AI courseware in school settings. First, instructional designs that incorporate AI courseware should move beyond activities focused solely on problem-solving and concept delivery and instead include opportunities for student interaction and collaborative inquiry. In particular, because different science units emphasize different inquiry skills – such as those addressed in the units *Light and Lenses*, *Uses of Electricity*, and *Seasonal Changes* – the application of AI courseware should be differentiated to align with the specific inquiry demands of each unit.

These findings suggest that AI courseware should not be implemented merely as a rapid problem-solving tool that emphasizes immediate correctness and automated feedback. Because many current AI courseware systems are primarily organized around individualized item-solving activities, students may have insufficient opportunities to engage in extended inquiry, reflective reasoning, and collaborative meaning-making processes. In particular, higher-order inquiry practices, such as generating investigable questions, evaluating alternative explanations, and constructing evidence-based arguments, require sufficient time for discussion, interpretation, and reflection beyond selecting correct answers. Therefore, deliberate instructional mediation by teachers is essential when integrating AI courseware into science classrooms.

Teachers should intentionally provide opportunities for students to pause, reflect, discuss, justify, and reinterpret their responses during AI-supported learning activities. For example, AI-generated tasks can be reorganized into small-group inquiry discussions or used as prompts for evidence-based reasoning and argumentative communication. In this respect, the educational effectiveness of AI courseware may depend not only on the technological functions of the system itself, but also on how teachers integrate AI-supported activities within inquiry-based and interaction-oriented instructional frameworks.

In addition, integrating real-life contexts with textbook-based inquiry activities can enhance students' engagement and promote more meaningful participation. To extend the positive effects observed in the numerical communication domain, students should have opportunities to engage in scientific communication using diverse representational forms, including written text, diagrams, and tables. This can be supported by the deliberate reconstruction by teachers of courseware tasks to encourage students to express their thinking in multiple ways.

Furthermore, teachers must maintain flexibility in reorganizing and adapting the instructional flow of AI courseware according to lesson objectives. Given that the courseware used in this study primarily emphasized individual problem-solving and automated feedback, active teacher mediation is essential to complement activities that require students to generate inquiry questions, design investigation procedures, and engage in interpretive discussions. Rather than using courseware tasks as isolated practice activities, teachers should integrate them into lessons that emphasize inquiry elements such as hypothesis generation, comparison of experimental designs, and collective interpretation of data.

This study has several limitations. Because only a single type of AI courseware was used, the findings cannot be generalized to all forms of AI courseware-based instruction. In addition, the sample was limited to a small number of classes from a specific region, and the intervention period was relatively short, making it difficult to examine long-term effects. Future research should explore diverse instructional designs and implementation strategies to better support inquiry-based science learning through AI courseware. In addition, continued field-based and practice-oriented studies are needed to ensure that AI courseware can contribute meaningfully to the development of scientific inquiry skills and scientific communication skills.

Beyond instructional implementation, the findings of this study also point to the need for the development of AI courseware that is explicitly designed to support scientific inquiry processes. The limited effects observed in inquiry and higher-order communication skills suggest that AI courseware focused primarily on individualized problem-solving and automated feedback may be insufficient to foster inquiry-oriented learning. Future AI courseware should therefore be designed to scaffold key inquiry practices, such as question generation, experimental planning, comparison of alternative

investigation designs, and evidence-based reasoning, in ways that promote student interaction and reflection. Importantly, such development should be grounded in close collaboration between educators and developers, ensuring that AI functionalities are aligned with curricular goals and pedagogical principles rather than operating as isolated technological solutions.

REFERENCES

- Ahn, J.S., & Kim, J.K. (2025). The effects of AI courseware-based problem-solving instruction on elementary students' academic emotions: Differences related to cooperative learning and learner characteristics. *Journal of Educational Technology*, 41(1), 205-237.
- Bray, A., & Tangney, B. (2017). Technology usage in mathematics education research: A systematic review of recent trends. *Computers and Education*, 114, 255-273.
- Campbell, D.T., & Stanley, J.C. (2015). *Experimental and Quasi-Experimental Designs for Research*. United States: Ravenio Books.
- Choi, I.H., & Sim, J.H. (2024). The effects of science classes using interactive digital textbooks on the self-directedness of elementary school students' and their creative problem-solving ability. *Biology Education*, 52(4), 700-712.
- Choi, S.Y., Cha, N.R., Jung, H.S., Noh, H.N., & Kim, D.K. (2024). Elementary teachers' perceptions of the adoption of AI digital textbooks. *Journal of Learner Centered Curriculum and Instruction*, 24(18), 769-791.
- Choi, S.Y., & Seo, J.H. (2009). The effect of using a digital science textbook on science problem-solving in elementary school students. *Journal of Korean Elementary Science Education*, 28(2), 132-141.
- Choppin, J., Carson, C., Borys, Z., Cerosaletti, C., & Gillis, R. (2014). A typology for analyzing digital curricula in mathematics education. *International Journal of Education in Mathematics Science and Technology*, 2(1), 11-25.
- Cohen, J. (2013). *Statistical Power Analysis for the Behavioral Sciences*. 2nd ed. London: Lawrence Erlbaum.
- Holmes, W., Bialik, M., & Fadel, C. (2019). *Artificial Intelligence in Education: Promises and Implications for Teaching and Learning*. United States: Center for Curriculum Redesign.
- Hoyles, C. (2018). Transforming the mathematical practices of students and teachers through digital technology. *Research in Mathematics Education*, 20(3), 209-228.
- Hoyles, C., & Lagrange, J.B. (Eds.). (2009). *Mathematics Education and Technology: Rethinking the Terrain New ICMI Study Series*. Vol. 13. Berlin: Springer.
- Jang, J.H., & Lee, E.J. (2024). Exploring teachers' acceptance levels and improvement strategies for AI-based teaching and learning platforms. *Educational Information and Media Research*, 30(6), 1659-1687.
- Jeon, S.S. (2019). Development and application of process-focused assessment to improve scientific communication skills. *Journal of Korean Elementary Science Education*, 38(1), 16-30.
- Joo, Y.J., & Lim, Y.J. (2015). Factors influencing student satisfaction with using digital textbooks in a middle school science class. *Journal of Curriculum and Instruction*, 19(2), 239-257.
- Kim, D., & Kim, W. (2025). A study on expanding the functions of the AI digital textbook (AIDT) to support inquiry-based learning. *The Korean Association of Information Education Research Journal*, 3(1), 83-90.
- Kwon, J.S., & Kim, B.K. (1994). Development of an instrument for the measurement of science process skills of Korean elementary and middle school students. *Journal of the Korean Association for Science Education*, 14(3), 251-264.
- Lee, J.H., & Gu, E.S. (2024). AI courseware-based education for lower-grade elementary school students with limited Korean literacy. *Journal of Creative Information Culture*, 10(4), 243-250.
- Lee, K.H., & Lee, J.H. (2024). A proposal to develop AI digital textbooks for elementary social studies with a focus on the geography section. *Journal of the Korean Association of Geographic and Environmental Education*, 32(3), 33-50.
- Lee, S.A. (2025). Effect of AI courseware usage on elementary students' academic self-efficacy. *Journal of the Korean Association of Information Education*, 29(1), 27-34.
- Lee, S.H., & Lee, Y.H. (2024). A study on elementary school teachers' perceptions of the introduction of AI digital textbooks. *Journal of the Korean Association of Information Education*, 28(4), 377-386.
- Lee, Y.A. (2024). Diagnostic and personalized reading instruction in elementary school English using AI-powered digital textbooks. *Korean Journal of Elementary Education*, 35(4), 73-91.
- Lemke, J.L. (1990). *Talking science: Language, learning, and values*. New York: Ablex Publishing Corporation.
- Lim, M.S. (2024). Development of an AI-based reading class model: Eye-pattern analysis and diagnostic courseware for sixth-grade students. *Journal of Reading Research*, 73, 113-140.
- Means, B., Bakia, M., & Murphy, R. (2014). *Learning Online: What Research Tells us about whether, when and how*. London: Routledge.
- Ministry of Education. (2022). *Elementary School Curriculum (Ministry of Education Notice No. 2022-33 [Booklet 2])*. India: Ministry of Education.
- Seo, S.S., Seo, J.H., & Hwang, S.H. (2009). Effects of the use of digital textbooks on improving problem-solving competence. *Journal of the Korean Association of Information Education*, 13(3), 263-271.
- Shin, J.H., & Hwang, Y.A. (2024). Development and validation of in developing sign guidelines for AI courseware-based classes in university general education. *Educational Information and Media Research*, 30(2), 571-594.
- Son, T.K., & Kang, D.H. (2024). Digital typological analysis of AI courseware in mathematics education. *Education of Primary School Mathematics*, 27(3), 261-279.
- Song, H.D., & Park, J.H. (2009). Effects of types of digital textbook use and levels of academic achievement on learning outcomes from an instructional guidance perspective. *Educational Information and Media Research*, 15(2), 29-46.
- Yoo, H.J., Yang, H.R., Oh, S.Y., Kim, T.E., Choi, J.I., & Park, S.E. (2025). Development of a digital supplementary textbook to enhance self-directed learning and motivation of elementary students: A focus on the plant unit of the elementary practical arts curriculum. *Journal of the Korea Institute of Information and Communication Engineering*, 29(6), 771-779.
- Zucker, T.A., Moody, A.K., & McKenna, M.C. (2009). The effects of electronic books on the literacy and language outcomes of pre-kindergarten to-grade 5 students: A research synthesis. *Journal of Educational Computing Research*, 40(1), 47-87.