International Journal of Evaluation and Research in Education (IJERE)
Vol. 15, No. 4, August 2026, pp. 3060~3074
ISSN: 2252-8822, DOI: 10.11591/ijere.v15i4.39254

Developing a 2-DOF robotic arm kit for embodied AI learning in middle school technology education

Dasol Kim[1], Sooin Kim[2]
[1]Chungnam Middle School, Daejeon, Republic of Korea
[2]CEO, Chang-ui-wa Gyoyuk, Gyeonggi-do, Republic of Korea

Article Info

Article history:

Received Feb 26, 2026
Revised Jun 24, 2026
Accepted Jul 6, 2026

Keywords:

AI literacy
Educational robotics
Physical computing
Robot-arm kit
Technology education

ABSTRACT

Artificial intelligence (AI) literacy is increasingly emphasized in technology education, yet many middle-school classroom activities remain screen-based and offer limited opportunities to experience the full AI pipeline in an authentic design context. To address this gap, this study developed a low-cost, classroom-ready 2-degree-of-freedom (2-DOF) robotic arm kit and an eight-lesson AI-integrated instructional unit in which students assembled the robot arm, collected and labeled image data, trained and improved a classification model, and applied model outputs to robot-arm control. A quasi-experimental pretest–posttest control-group design was employed with 98 ninth-grade students in four intact classes (experimental group, n=46; comparison group, n=52). Quantitative data were analyzed using ANCOVA with pretest scores as covariates, and qualitative data from student reflection reports were analyzed thematically. The experimental group significantly outperformed the comparison group on value of AI, efficacy of AI, AI literacy, and total scores. Qualitative findings further showed that students recognized both benefits and risks of AI and proposed multi-layered safety strategies involving physical safeguards, operational rules, and data/model management. These findings suggest that embedding the full AI pipeline in a tangible engineering-design task can support embodied, responsibility-oriented AI learning in middle school technology education.

Corresponding Author:

Dasol Kim
Chungnam Middle School
123, Gyejok-ro, Dong-gu, Daejeon 34641, Republic of Korea
Email: losad721@naver.com

1. INTRODUCTION

Artificial intelligence (AI) is rapidly being embedded not only in everyday digital services such as search, recommendation, and translation, but also in school administration, learning platforms, and assessment tools. Consequently, adolescents are increasingly expected to develop AI literacy that extends beyond merely “using” AI—namely, the capacity to interpret and evaluate how AI systems operate (e.g., what data and rules they rely on), to recognize their limitations, biases, and risks, and to use them responsibly within real-world contexts [1], [2]. The international community likewise highlights the importance of K–12 AI education and recommends that educational policies address AI competencies together with ethics and accountability [1]–[4]. In particular, the view that trustworthy AI requires systematic management of key risk factors—such as fairness, safety, transparency, and accountability—has become an important benchmark for educational practice; accordingly, the National Institute of Standards and Technology (NIST) AI Risk Management Framework provides guidance for identifying, assessing, and mitigating risks across the AI lifecycle [5].

AI literacy also offers a basis for specifying what should be taught and what should be assessed. Long and Magerko [6] articulated core AI literacy competencies (e.g., understanding AI concepts; understanding relationships among data, models, and outputs; critically judging limitations, errors, and societal impacts) and provided design considerations for instruction. Ng *et al.* [7] further conceptualized AI literacy along dimensions such as understanding, use, evaluation/creation, and ethics, strengthening the theoretical foundation for curriculum and assessment design. However, the implementation of K–12 AI education remains uneven relative to policy aspirations. Touretzky *et al.* [8] proposed the scope and principles of AI learning that all K–12 students should acquire and emphasized learning progressions organized around the "big ideas" of AI, including machine learning [9]. Despite these proposals, recent synthesis studies report that classroom AI instruction often prioritizes conceptual introductions and tool use, whereas learning experiences that consistently support higher-order competencies—such as learners constructing datasets, iteratively improving models, and interpreting performance and limitations—remain comparatively scarce [10], [11]. Large-scale analyses of classroom instruction similarly indicate substantial variation in learning theories, pedagogical approaches, and assessment alignment, and suggest that performance-based verification of AI literacy outcomes is still limited [12].

To demonstrate "what has improved," reliable measurement instruments are essential. Work developing and validating AI literacy scales for middle school students proposes multidimensional constructs—including AI understanding, data literacy, problem solving, and social impact/ethics—and indicates the feasibility of school-based diagnostics [13]. In addition, general AI literacy scale development provides a basis for cross-context comparison and empirical validation of learning effects [14]. Importantly, ethics and social impact should not be treated as optional add-ons in AI education but as core components. Scholarship has increasingly called for integrating socio-ethical issues into K–12 AI learning experiences [15], and studies on project-based, integrated AI-and-ethics curricula suggest that activity-centered learning, continuous embedding of ethical reasoning, and reduced access barriers can enhance engagement and understanding [16].

This study operationalizes these discussions within technology education, which is grounded in engineering design and maker-oriented learning. Technology education connects knowledge and practice by guiding learners through defining problems, specifying designs, building and testing prototypes, and providing evidence-based rationales for improvement. The project-based learning (PBL) literature emphasizes that autonomy and authenticity alone are insufficient; rather, structured inquiry support (scaffolding) and assessments aligned with learning goals are decisive conditions for learning outcomes [17], [18]. Moreover, research on scaffolding in problem-/project-based learning argues that effective instruction requires deliberately balancing cognitive challenge with instructional support, rather than relying on "minimal intervention" [19], [20]. Therefore, integrating AI learning into technology-education projects calls for explicitly embedding the AI pipeline (data collection → training → evaluation → application) within the activity flow and making learners' decision rationales visible at each stage (e.g., data quality, class balance, error types, uncertainty).

Educational robotics and physical computing are particularly promising because they make AI decisions observable as physical outcomes, thereby visualizing how model outputs affect real-world actions. Systematic reviews of robotics-based learning suggest that robotics can enhance motivation and achievement, although effects vary depending on task design, teacher scaffolding, and assessment practices [21]. Reviews of robots in education likewise note that learning outcomes depend on the role of the robot (tool, tutor, peer), instructional context, and reliability issues [22]. In addition, low-cost microcontroller-based instruction (e.g., Arduino) can promote programming and computational thinking in PBL/STEM contexts [23], [24], and prior work on low-cost microcontroller-based teaching models demonstrates feasibility and scalability for school implementation [25]. Block-based languages (e.g., Scratch) lower entry barriers for novice learners [26], and the growing availability of web-based interfaces for training and testing classification models has further enabled machine-learning activities in K–12 contexts [27].

Nevertheless, in Korea, middle school AI education is primarily implemented within informatics/computing subjects, and prior analyses have raised concerns about weak vertical articulation across school levels due to limited instructional time and variability in curriculum organization [28]. Even when physical computing is used, instruction often concentrates on rule-based programming (e.g., conditional branching from sensor readings), which may restrict opportunities for learners to construct datasets, evaluate and improve model performance, and mitigate the consequences of misclassification [10]–[12]. Although fusion education attempts using 3D printing for robot-arm fabrication and control have been reported [29], classroom-ready designs that reduce burdens related to assembly, circuit configuration, and lesson operation at the middle school level remain an ongoing challenge. Meanwhile, AI-based invention education programs have demonstrated the potential of AI for creative problem solving and design activities [30]. However, concrete instructional packages that enable middle school students to experience the integrated flow of model outputs → physical actions → responsible design (e.g., safety, reliability, misuse) require further development and empirical validation.

Accordingly, this study aims to i) develop a low-cost, classroom-ready 2-degree-of-freedom (2-DOF) robot-arm kit that middle school students can assemble with relative ease while learning basic mechanical elements and simple electronic circuits; ii) support students in training, evaluating, and improving an image-classification model using beginner-friendly web-based tools and a block-coding environment; and iii) design, validate, and evaluate an AI-integrated instructional unit that implements the AI pipeline (data → model → evaluation → application) within an engineering design task by linking model outputs to robot-arm movements. The 2-DOF configuration was deliberately selected because it allows accessible kinematic understanding while still enabling learners to experience engineering constraints such as workspace limits, load constraints, and motion accuracy. At the same time, the kit was intentionally designed for ordinary school conditions by reducing assembly complexity, simplifying control logic, and optimizing the system for USB single-cable operation without a separate external power supply. In this way, the study sought to develop not only an instructional tool, but also a feasible embodied AI learning platform that could be implemented even in resource-constrained school settings. This design is also meaningful in that it enables discussion of AI risk management and responsible design by allowing students to observe how misclassification may lead to physical misoperation [5], [21], [22].

This study is novel in four respects. First, unlike many educational robot-arm activities that rely on more complex multi-joint systems, external power supplies, or screen-based AI exploration, this study intentionally adopts a 2-DOF structure and a USB single-cable design to maximize accessibility, operational stability, and classroom feasibility under ordinary school conditions. This design reduces assembly difficulty, simplifies control logic, and enables embodied AI learning even in under-resourced school settings, including classrooms with older PCs and limited infrastructure. Second, the study does not treat the robot arm merely as a programmable artifact; rather, it presents a power-optimized, school-scalable instructional kit designed so that novice learners can assemble, operate, and test an embodied AI system within realistic class constraints. Third, the instructional unit embeds the full AI pipeline—data collection, labeling, model training, evaluation, revision, and deployment—within a tangible engineering-design task, allowing students to observe how model outputs and model errors translate into physical action and potential misoperation. Fourth, by combining a quasi-experimental design with qualitative reflection data, this study examines not only whether students' AI-related outcomes improve, but also how embodied AI learning can promote reasoning about safety, misuse, and responsible AI governance in middle school technology education. In this respect, the study contributes not only to embodied AI learning design but also to the broader dissemination of robotics and AI education by proposing a kit that is feasible for routine use in resource-constrained school environments. The research questions are as:

- Are the developed 2-DOF robotic arm kit and the accompanying instructional materials feasible and appropriate for implementation in middle school technology education? (RQ1)
- Compared with conventional technology education, does participation in the unit improve students' AI-related outcomes (value, efficacy, and AI literacy)? (RQ2)

2. PROCEDURE FOR DEVELOPING THE INSTRUCTIONAL KIT AND MATERIALS

2.1. Overview of kit design

To develop the instructional kit, we drew on a previously reported low-cost microcontroller-based teaching and learning model [25]. Following the procedure shown in Figure 1: i) we selected the topic and determined that a 2-DOF robot arm with a gripper would be appropriate for practicing AI image-classification models; ii) fabricated an Medium-density fiberboard (MDF)-based kit using an Arduino-compatible board; iii) conducted pilot tests to confirm classroom applicability; iv) produced instructional materials using Canva; v) we implemented the kit and unit with middle school students to examine educational effects; and vi) refined the kit and materials based on identified areas for improvement. This process was intended to ensure that the kit was not only technically functional but also instructionally manageable in an ordinary middle-school setting.

The kit consists of a two-servo-joint 2-DOF robot-arm mechanism, a microcontroller-based control module, a gripper, and a manually operated joystick. To facilitate adoption in typical middle school classrooms, the mechanical components were designed using low-cost materials, primarily MDF and an Arduino-compatible board. The kit was also designed to reduce assembly burden, simplify circuit connections, and support repeated classroom use under limited time and resource conditions.

2.2. AI learning flow

The instructional sequence was designed to enact the core AI pipeline emphasized in K–12 AI frameworks [6], [9]. Students first defined a real-world problem for which classification would be useful, such as sorting objects by category or recognizing simple hand gestures. They then collected and labeled training data and trained a classification model using Entry, a beginner-friendly web-based tool [27].

Students interpreted the confidence levels of image-classification outputs by examining the collected dataset and the resulting model behavior. They iteratively improved model performance by adjusting the quantity and quality of data, including class balance, diversity of examples, lighting, and background conditions. Finally, students deployed the trained model by connecting it to robot-arm control, observed how misclassification could lead to physical misoperation, and discussed mitigation strategies such as adding an "unknown/no-action" class, using repeated predictions, improving data quality, or redesigning the task. In this way, the unit allowed students to experience AI not only as software output, but also as a decision process with visible physical consequences.

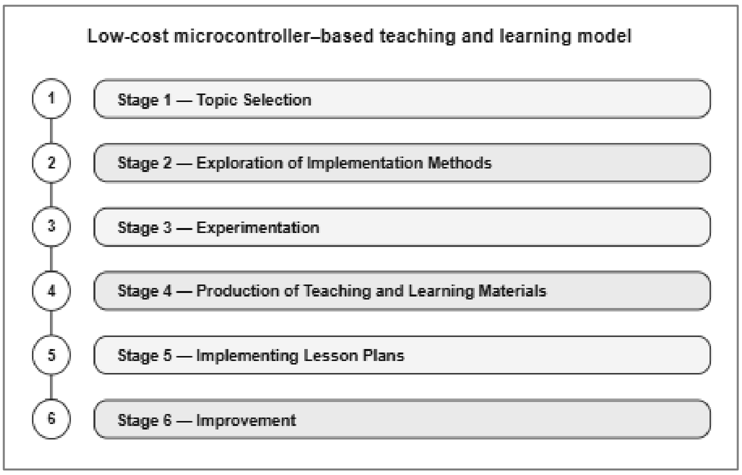

Figure 1. Overview of kit design

2.3. Unit structure

The unit was designed as eight 45-minute lessons. It combined brief concept instruction, guided practice, team-based design work, and reflection, aligning with recommended practices for implementing PBL. The lesson overview is presented in Table 1.

Table 1. Unit overview

Lesson	Core topic	Key deliverables
1	Problem definition and safety	Problem statement; list of constraints
2	Robot-arm assembly and basic control	Operating robot arm; movement log/record
3	Classification concepts and data understanding	Class definitions; data collection plan
4	Data collection and labeling	Training dataset v1
5	Model training and evaluation	Model v1 metrics; error-analysis notes
6	Model improvement and fairness/limitations	Model v2; mitigation strategy plan
7	Application: integrating model output with robot control	Demonstration results; failure-cause analysis
8	Assessment and reflection	Design report; responsible AI reflection essay

3. METHOD

3.1. Research design

This study adopted a quasi-experimental pretest–posttest control-group design and a convergent mixed-methods design. The quantitative component examined whether participation in an AI-integrated instructional unit based on a low-cost 2-DOF robotic arm improved students' AI-related outcomes compared with conventional technology education. The qualitative component was included to capture how students interpreted the value, risks, and responsible use of AI when model outputs were connected to a physical robotic system.

The experimental group participated in the AI-integrated unit developed in this study, whereas the comparison group completed an alternative eight-lesson sequence focused on robotics/mechanical elements and rule-based programming without machine-learning model training or application. This design was selected because random assignment at the individual student level was not feasible under ordinary school scheduling constraints. Quantitative and qualitative data were collected during the same intervention period, analyzed separately, and integrated during the interpretation stage to provide a more comprehensive explanation of both learning outcomes and students' reflective reasoning.

3.2. Participants and context

Participants were ninth-grade students from a public middle school in the Republic of Korea. A total of 98 students participated in the study, comprising two experimental classes (n=46) and two comparison classes (n=52). Exact age data were not collected in order to minimize the collection of personal information in this school-based study; however, ninth-grade students in the Korean school system are typically 15 to 16 years old. The experimental group included 26 male students and 20 female students, whereas the comparison group included 31 male students and 21 female students. All participating classes belonged to the same grade level and shared a similar school context in terms of curriculum exposure, timetable structure, and instructional environment.

Class assignment was conducted on an intact-class basis according to school timetable constraints between April 7 and April 25, 2025. Because the school schedule did not permit random reassignment of students across classes, the study should be interpreted as a school-based quasi-experimental investigation rather than as a randomized controlled trial. Within classes, students worked in teams of three to four, and team formation was conducted through random assignment.

Both groups were taught by the same technology teacher, which helped reduce teacher-related variation across conditions. Prior to implementation, approval was obtained from the school principal as the relevant institutional authority, and both parental consent and student assent were secured. Participation was voluntary, and students were informed that their responses would not affect their grades. Although the intact-class design increased ecological validity by preserving authentic classroom conditions, the findings should still be interpreted with caution because class-level clustering effects may remain.

3.3. Treatment and comparison conditions

3.3.1. Experimental group

Students in the experimental group completed the eight-lesson AI-integrated unit. The intervention consisted of eight 45-minute lessons, for a total of 360 minutes, implemented over the scheduled class period. Across the lessons, students defined a classification-related problem, assembled and controlled a 2-DOF robotic arm, collected and labeled image data, trained and evaluated an image-classification model using Entry, improved the model through iterative data revision, and finally connected model outputs to robot-arm movement.

3.3.2. Comparison group

Students in the comparison group completed an eight-lesson conventional technology education sequence, also totaling 360 minutes. This sequence focused on robot-arm assembly, servo control, and rule-based programming. Although the comparison group worked on a similar physical artifact and completed a similar design-related task, their lessons did not involve machine-learning model training, model evaluation, or model-based control. Thus, the key instructional difference between the two conditions was the inclusion or exclusion of machine-learning activities and AI-pipeline reasoning.

3.3.3. Implementation monitoring

To monitor implementation fidelity, the instructor used a fidelity checklist that recorded the completion of key lesson activities, time allocation, and delivery of core concepts. Example items included whether planned safety instruction was delivered, whether model training and testing steps were completed as scheduled, and whether teams produced the required artifacts for each lesson. In addition, students' activity status and portfolio progress were reviewed through shared access to their Google Sites portfolios. These procedures were used to confirm that the intended instructional sequence had been implemented consistently across the intervention period.

3.4. Instruments

3.4.1. Quantitative measures

To assess AI-related outcomes, this study adopted an AI literacy instrument reported in prior Korean educational research. The instrument comprised 14 items across three subdomains: value of AI (5 items), efficacy of AI (4 items), and AI literacy (5 items). All items were rated on a 5-point Likert scale ranging from 1 (strongly disagree) to 5 (strongly agree). Representative items addressed whether students regarded AI as socially useful and important, whether they felt confident in understanding and using AI-related ideas, and whether they could explain how AI systems work and why AI outputs may sometimes be inaccurate or biased.

Before the main implementation, the instrument was pilot-tested in a separate class (n=26) to examine item clarity and suitability for the target age group. Based on pilot feedback, minor wording revisions were made to improve readability and reduce ambiguity. Internal consistency reliability in the present dataset was high, with Cronbach's alpha values ranging from 0.858 to 0.959 across the measured variables. However, the dimensional structure of the instrument was not revalidated through factor analysis in the present sample, so construct validity in this specific context should be interpreted with caution.

3.4.2. Qualitative data sources

Qualitative data were collected to complement the survey findings and to provide a richer account of how students reasoned about AI in a physical computing context. Students documented the production process, models, and code in their design reports. In addition, students responded to four open-ended prompts: i) useful tasks that could be done with an AI robot arm; ii) expected benefits of advances in AI; iii) concerns regarding advances in AI; and iv) rules for using AI safely.

These prompts were designed to elicit students' perceptions of applicability, benefits, risks, and responsible use. The written responses were supplemented by related artifacts such as dataset screenshots, model-evaluation records, and portfolio materials. Together, these sources were used as the primary qualitative dataset for analysis and interpretation.

3.5. Data collection procedures

The pretest was administered one week prior to the beginning of the instructional unit. The posttest was administered within one week after the intervention had concluded. In addition to the survey data, student artifacts—including datasets, model-evaluation screenshots, design reports, and portfolio materials—were collected in an open-portfolio format using Google Sites. Before analysis, all data were anonymized by replacing student names with participant codes, and the multiple data sources were later integrated during interpretation.

3.6. Data analysis

3.6.1. Quantitative analysis

Quantitative analyses were conducted in several steps. First, the data were screened and descriptive statistics were computed for each measure. Second, baseline equivalence between the experimental and comparison groups was examined using independent-samples t-tests on pretest scores. Third, treatment effects were tested using separate analyses of covariance (ANCOVAs) for each outcome variable, with the corresponding posttest score as the dependent variable, group (experimental vs. comparison) as the fixed factor, and the corresponding pretest score as the covariate.

Prior to the main analyses, assumption checks were conducted. For each ANCOVA, the assumption of homogeneity of regression slopes was tested by examining the Group × Pretest interaction. In addition, normality was examined through Shapiro–Wilk tests and visual inspection of Q–Q plots, and homogeneity of variance was examined using Levene's test. These checks indicated some deviations from normality and homogeneity of variance for several posttest outcomes, likely reflecting ceiling effects in a portion of the data. Because of these deviations, the ANCOVA results were interpreted with caution and were supplemented by robustness checks using heteroskedasticity-robust standard errors.

3.6.2. Qualitative analysis

The qualitative data were analyzed using thematic analysis following Braun and Clarke's six-phase approach [31]. First, the researcher repeatedly read students' written responses and artifacts to gain familiarity with the data. Second, initial meaning units were identified and coded. Third, related codes were grouped into categories, which were then reviewed and refined into broader themes. Finally, the themes were defined, named, and interpreted in relation to the research purpose and quantitative findings.

The coding process followed a hybrid approach. The four reflection prompts provided an initial deductive frame for organizing the analysis, while specific categories and subthemes were generated inductively from students' actual responses. In the current dataset, this process yielded 226 meaning units and four overarching themes: i) perceived applicability; ii) expected benefits; iii) concerns and risks; and iv) rules for safe and responsible use. To enhance traceability, coding proceeded from raw statements to meaning units, codes, subthemes, and final themes. For example, a student response such as "The robot should stop immediately if something goes wrong" was coded as emergency stopping and then categorized under the subtheme of operation control within the broader theme of rules for safe and responsible use.

3.7. Research ethics

This study was conducted in accordance with school policies and relevant regulations governing school-based educational research. Formal approval for the study was obtained from the school principal, who served as the relevant institutional authority for this school-level implementation. In addition, both parental consent and student assent were secured prior to data collection.

Participation was voluntary and had no effect on students' grades. To minimize risks associated with the robot arm, the researcher personally provided safety instruction and supervised student use under the guidance of the classroom teacher. To reduce possible educational inequity, supplementary instruction related to the AI-integrated activities was provided to the comparison group after the study had concluded.

4. RESULTS AND DISCUSSION

4.1. Development of the 2-DOF robot-arm kit

The robot-arm kit was designed to be accessible for middle school learners and affordable for implementation in general public schools with limited budgets. To reduce fabrication costs while maintaining structural robustness, the mechanical frame was constructed using two MDF plates, and the control system was built on an Arduino-compatible board. To simplify wiring and lower barriers for classroom use, an Arduino sensor shield was adopted for circuit connections, and a manually operated joystick was added to support basic control and testing.

To examine the appropriateness of the kit for the intended learners and classroom context, the developed prototype underwent expert review by eight evaluators: five faculty members/researchers specializing in technology education, engineering education, and AI education, and three in-service technology teachers. The review was designed to assess the feasibility and appropriateness of both the kit and the accompanying instructional materials for middle school implementation. Specifically, experts provided structured feedback on i) classroom usability (e.g., assembly difficulty, operability, and manageability); ii) operational stability and safety (e.g., power stability and potential risk points); and iii) instructional suitability (e.g., alignment with learning objectives and age appropriateness). In addition, implementation notes from the pilot lessons (e.g., setup time, common operational issues, and safety-related observations) were collected to support a practice-oriented feasibility assessment. Overall, the panel judged the kit to be appropriate for middle school students and likely to support engaging learning tasks. The kit components and the completed prototype are presented in Figure 2.

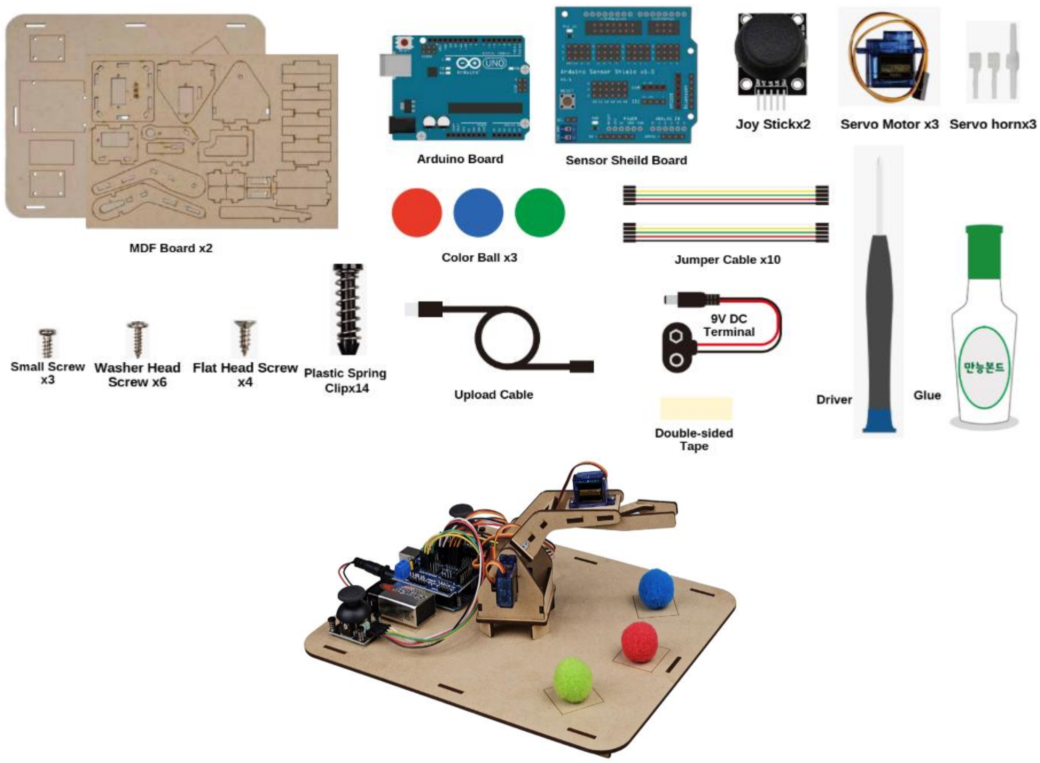

Figure 2. Components and build quality of the 2-DOF robot-arm kit

This study intentionally adopted a 2-DOF robotic arm configuration rather than a conventional 3-degree-of-freedom (3-DOF) educational kit to maximize classroom usability and operational stability. First, reducing the number of joints and components improves ease of use: students with limited prior experience can assemble the mechanism more quickly, and the corresponding control logic becomes simpler to implement and debug. This simplification supports instructional goals by lowering initial barriers to entry and allowing learners to focus on core concepts such as sensing–decision–actuation and error diagnosis.

Second, the kit was designed to operate with a "USB single-cable" setup, where one USB connection to the host computer provides both communication and power, eliminating the need for a separate external power supply. This decision was motivated by typical school conditions (e.g., older classroom PCs and limited outlets), where stable operation is more likely when the entire system remains within conservative USB power limits. In general, USB 2.0 ports provide up to approximately 500 mA and USB 3.0 ports up to approximately 900 mA under nominal conditions; therefore, the design goal was to complete all operations within the lower bound (≤500 mA) to reduce the risk of power-reset events and ensure uninterrupted practice.

To justify this constraint, the current-consumption characteristics of SG90 servo motors were considered during both mechanical and instructional design. Under typical motion, an SG90 servo may draw roughly 100–300 mA; however, current draw increases when the servo must hold a load (holding current) or when it becomes mechanically constrained (stall condition), which can reach approximately 600–800 mA for a single servo—potentially exceeding the USB 2.0 threshold by itself. Because the gripper servo is particularly susceptible to high load during grasping and holding, accurate neutral-position calibration and careful mechanical alignment were emphasized, and a lightweight compliant object (a sponge ball) was selected as the primary target to reduce stall risk and prevent overload.

Third, additional strategies were implemented to enhance safety and reliability under strict power constraints. The control routine was structured so that servos do not move simultaneously; instead, they operate sequentially to distribute peak current demand over time. Although some combined states inevitably occur (e.g., one servo moves while another maintains holding torque), sequential actuation reduces instantaneous current spikes. At the hardware level, the use of a sponge ball rather than rigid objects further mitigates stall current at the gripper. Finally, component optimization—especially servo zeroing and linkage-angle tuning—was used to minimize unnecessary mechanical resistance, thereby improving both stability and repeatability of classroom operation.

4.2. Development of instructional materials

Instructional materials were developed to enable students to carry out hands-on activities with minimal difficulty and to support teachers' implementation fidelity. The materials included: i) a component checklist; ii) step-by-step guidance for building an open portfolio using Google Sites; iii) robot-arm assembly instructions; iv) procedures for creating an AI image-classification model using Entry, including example code for basic practice; v) joystick-control instructions prepared for the comparison condition; and vi) templates and guidance for producing reflection artifacts. These materials were designed not only to assist students' task completion but also to reduce teachers' preparation burden and improve lesson consistency across groups.

All resources and student portfolios were shared via Padlet, allowing teachers to quickly access instructional materials and enabling teams to monitor progress across groups during the unit. This sharing system also supported formative monitoring by making students' work processes visible during implementation. The developed materials are shown in Figure 3.

Figure 3. Developed instructional materials

4.3. Analysis of educational effects
4.3.1. Baseline equivalence between groups

Independent-samples t-tests on pretest scores indicated that the experimental and comparison groups were statistically equivalent at baseline on value of AI, efficacy of AI, literacy of AI, and the total score (value: $t(96)=.176$, $p=.861$; efficacy: $t(96)=-1.535$, $p=.128$; literacy: $t(96)=-.752$, $p=.454$; total: $t(96)=-.718$, $p=.474$). Cronbach's alpha values ranged from .858 to .959, indicating high internal consistency reliability. The pretest descriptive statistics by group are summarized in Table 2.

Table 2. Baseline equivalence between groups at pretest

Type	Group type	N	Mean	SD	t(df)	p	Cronbach's α
Value of AI	Experiment	46	3.60	.953	.18(96)	.861	.858
	Comparison	52	3.58	.565			
Efficacy of AI	Experiment	46	3.44	1.087	-1.54(96)	.128	.917
	Comparison	52	3.70	.542			
Literacy of AI	Experiment	46	3.61	1.107	-.75(96)	.454	.938
	Comparison	52	3.74	0.514			
Total score	Experiment	46	3.56	0.990	-.72(96)	.474	.959
	Comparison	52	3.67	0.501			

An ANCOVA was conducted for each outcome to compare posttest scores between the experimental (E) and comparison (C) groups, controlling for the corresponding pretest scores. The homogeneity-of-regression-slopes assumption was satisfied for all models, with non-significant Group × Pretest interactions. However, normality and homogeneity-of-variance checks indicated some deviations for several posttest variables, likely due to ceiling effects. Accordingly, the ANCOVA results were interpreted together with robustness checks using heteroskedasticity-robust standard errors, and the overall pattern of significant group differences remained unchanged. As shown in Table 3, significant adjusted group effects were found for all outcomes, favoring the experimental group.

Table 3. Posttest group comparisons adjusted for pretest (ANCOVA)

Type	F(1,95)	p (Holm)	partial η^2	Adj M (Exp, E)	Adj M (Comp, C)	Δ (C−E)
Value of AI	6.720	.011	.066	3.986	3.639	−0.347
Efficacy of AI	10.331	.007	.098	4.149	3.680	−0.469
Literacy of AI	10.393	.007	.099	4.185	3.715	−0.470
Total Score	10.283	.007	.098	4.103	3.684	−0.419

4.3.2. Pre–post changes in the experimental group

Following participation in the 2-DOF robot-arm–based AI-integrated unit, the experimental group showed statistically significant improvements across AI-related perceptions and competencies. Value of AI increased from pretest (M=3.60, SD=.953) to posttest (M=3.99, SD=.842), (p=.013). Efficacy of AI increased from M=3.44 (SD=1.087) to M=4.09 (SD=.978), (p<.001). Literacy of AI increased from M=3.61 (SD=1.107) to M=4.16 (SD=.981), (p=.002). The total score increased from M=3.56 (SD=.990) to M=4.08 (SD=.872), (p<.001). These results suggest that the intervention was associated with meaningful gains in students' AI-related perceptions and AI literacy outcomes. The magnitude of within-group change in the experimental group was small-to-moderate for value of AI (dz=0.38) and moderate for efficacy of AI (dz=0.60), AI literacy (dz=0.48), and total score (dz=0.52). Detailed pre–post results for the experimental group are provided in Table 4.

Table 4. Pre–post changes in the experimental group

Type	test	N	Mean	SD	t(df)	p	Cohen's dz
Value of AI	Pre-test	46	3.60	.953	-2.59(45)	.013	.38
	Post-test	46	3.99	.842			
Efficacy of AI	Pre-test	46	3.44	1.087	-4.07(45)	<.001	.60
	Post-test	46	4.09	.978			
Literacy of AI	Pre-test	46	3.61	1.107	-3.23(45)	.002	.48
	Post-test	46	4.16	.981			
Total score	Pre-test	46	3.56	.990	-3.56(45)	<.001	.52
	Post-test	46	4.08	.872			

4.3.3. Pre–post changes in the comparison group

In contrast, the comparison group did not exhibit statistically significant changes on any outcome measure. Value of AI increased slightly from pretest (*M*=3.58, *SD*=.565) to posttest (*M*=3.63, *SD*=.568), but the change was not significant (*p*=.510). Efficacy of AI showed a small, non-significant change from *M*=3.70 (*SD*=.542) to *M*=3.73 (*SD*=.571), (*p*=.734). Literacy of AI remained stable (*M*=3.74 at both pretest and posttest; *p*=.994). The total score also showed no significant improvement, changing from *M*=3.67 (*SD*=.501) to *M*=3.70 (*SD*=.522), (*p*=.660). In contrast, the corresponding within-group effect sizes in the comparison group were negligible, ranging from *dz*=.00 to 0.09. The comparison-group results are presented in Table 5.

Table 5. Pre–post changes in the comparison group

Type	test	N	Mean	SD	t(df)	p	Cohen's dz
Value of AI	Pre-test	52	3.58	.565	-.66(51)	.510	.09
	Post-test	52	3.63	.568			
Efficacy of AI	Pre-test	52	3.70	.542	-.34(51)	.734	.05
	Post-test	52	3.73	.571			
Literacy of AI	Pre-test	52	3.74	.514	-.01(51)	.994	.00
	Post-test	52	3.74	.528			
Total Score	Pre-test	52	3.67	.501	-.44(51)	.660	.06
	Post-test	52	3.70	.522			

4.3.4. Qualitative analysis of reflection reports

To complement the quantitative findings, reflection reports from students in the experimental group were analyzed qualitatively. As shown in Table 6, the analysis of 226 meaning units yielded four overarching themes: i) perceived applicability; ii) expected benefits; iii) concerns and risks; and iv) rules for safe and responsible use. These themes provide insight into how students interpreted the educational and social meaning of AI when model outputs were connected to a physical robotic system.

Table 6. Themes and subthemes identified from students' reflection reports (n=226 meaning units)

Higher-order category	Subtheme (Frequency)	Representative statement (Excerpt/example)
Perceived applicability	Classification & organizing (16)	"Separating and inspecting defective products…," "Sorting ballot papers…"
	Recycling & environment (5)	"Replacing waste sorting with a robot arm…"
	Automation in manufacturing/logistics/services (6)	"Automatically sorting parts and supplying them to an assembly line."
	Hazardous/extreme environments (5)	"Sorting radioactive materials/dangerous waste… even in space or on cliffs."
	Healthcare & caregiving (4)	"It could be used as a precise surgical tool."
Expected benefits	Quality of life & convenience (12)	"Like a personal assistant… a more comfortable life."
	Risk reduction & safety (9)	"Supporting people working in harsh environments."
	Efficiency & productivity (5)	"Doing repetitive tasks quickly… continuous operation is possible."
	Cost & labor reduction (4)	"Reducing labor costs… helping to lower prices."
Concerns/risks	Job displacement (6)	"Jobs may be threatened… social problems could occur."
	Misuse (weaponization/surveillance) (6)	"Misused by criminal groups…," "Concerns about unauthorized surveillance."
	Malfunction & harm (3)	"If it malfunctions, production lines could stop… causing economic losses."
	Cost & maintenance burden (3)	"Installation and maintenance costs… difficult to use widely."
	Performance limitations (1)	"Hard to distinguish detailed materials in recycling (PP, PS)."
Safety rules	Emergency stop & operation control (13)	"Establish an emergency-stop system…," "No manual intervention during operation."
	Inspection & maintenance (9)	"Regular inspections and repairs… preventing electrical accidents."
	Safety zones & restricted access (7)	"Set safety fences/sensor boundaries… restrict access."
	Policy/training/certification (6)	"Only qualified users… national-level inspection/authorization."
	Security & data management (7)	"Prevent viruses…," "Update training data."

First, under perceived applicability, students most frequently described the robot arm in terms of classification, sorting, and organizing functions. They extended these functions to a variety of real-world contexts, including recycling, factory inspection, logistics, restaurant work, hazardous environments, and healthcare. The most frequent subtheme was classification and organizing (n=16), followed by automation in manufacturing/logistics/services (n=6), recycling and environment (n=5), hazardous or extreme environments (n=5), and healthcare and caregiving (n=4). These responses suggest that students were able to transfer the robot arm's core logic to broader application domains.

Second, under expected benefits, students emphasized convenience, efficiency, safety, and improved quality of life. The most frequent subtheme was quality of life and convenience (n=12), followed by risk reduction and safety (n=9), efficiency and productivity (n=5), and cost and labor reduction (n=4). This pattern indicates that students perceived AI not merely as a technical tool, but as a means of improving working conditions and supporting everyday life.

Third, under concerns and risks, students expressed awareness of both technical and social problems associated with AI applications. The most salient concerns were job displacement (n=6) and misuse, including weaponization or surveillance (n=6). Additional concerns included malfunction and harm (n=3), cost and maintenance burden (n=3), and performance limitations (n=1). These reflections indicate that students were not only enthusiastic about AI applications but also capable of identifying possible unintended consequences and real-world constraints.

Fourth, under rules for safe and responsible use, students proposed multi-layered safety strategies that extended beyond immediate hardware precautions. The most frequent subtheme was emergency stop and operation control (n=13), followed by inspection and maintenance (n=9), safety zones and restricted access (n=7), security and data management (n=7), and policy/training/certification (n=6). Notably, some students suggested not only physical safeguards but also data/model management practices, such as updating training data and checking system accuracy. This suggests an emerging understanding that safe AI use depends on both mechanical protection and ongoing AI-system management.

Overall, the qualitative findings support the quantitative results by showing that the intervention promoted not only more positive perceptions of AI but also more reflective reasoning about limitations, misuse, safety, and responsible governance. In particular, the qualitative themes help explain why the experimental group showed stronger gains than the comparison group, because students in the AI-integrated condition engaged not only in building and control but also in evaluating model behavior, anticipating failure, and proposing safeguards. Thus, the qualitative data strengthen the interpretation that the educational value of the intervention lay in embedding the full AI pipeline in a physical design task rather than in robotics activity alone.

4.4. Discussion

The findings of this study suggest that embedding the full AI pipeline within a tangible engineering-design task can meaningfully improve middle school students' value of AI, efficacy of AI, and AI literacy. Compared with the comparison group, the experimental group showed significantly higher posttest outcomes across all measured variables after controlling for pretest scores, and the experimental group also showed significant pre–post gains whereas the comparison group did not. Taken together, these results indicate that the educational effect was not simply due to participation in a robotics activity itself, but rather to students' engagement in core machine-learning processes such as data collection, labeling, model training, evaluation, revision, and deployment. This interpretation is consistent with AI literacy perspectives that emphasize understanding the relationships among data, models, and outputs rather than merely using AI tools at a surface level [6], [7], [9].

A key implication of the present findings is that the instructional unit helped move AI learning from an abstract, screen-based experience toward a more embodied and interpretable form of learning. By linking classification outputs to the visible motion of a robot arm, students could directly observe how data quality, class balance, labeling decisions, and model limitations influenced physical action. In this sense, the intervention made the relationships among data, model, and output more concrete and meaningful. This is important because recent reviews have pointed out that K–12 AI education often remains limited to conceptual introductions or tool use, with fewer opportunities for learners to iteratively build datasets, improve models, and interpret performance and limitations in context [10]–[12].

The gains in AI literacy are especially meaningful when interpreted in relation to existing conceptualizations of AI literacy. AI literacy should include not only conceptual understanding of AI, but also the ability to critically judge limitations, errors, and societal implications [6]. Similarly, Ng *et al.* [7] highlighted that AI literacy involves understanding, use, evaluation, and ethics as interconnected dimensions. The present instructional unit appears to support these dimensions by requiring students to train a model, interpret confidence levels, compare model versions, identify likely causes of failure, and propose mitigation strategies such as improving data quality or adding an "unknown/no-action" category. These practices go beyond isolated tool use and align with the argument that AI literacy should involve informed judgment about how AI systems function and where they may fail [6], [7], [9].

The increase in value of AI may be interpreted as reflecting a broader and more situated understanding of AI in social and everyday contexts. Students did not merely report that AI was useful in a general sense; they identified concrete applications in recycling, manufacturing, healthcare, logistics, and hazardous environments, while also recognizing concerns such as misuse, surveillance, and job displacement.

This pattern suggests that the instructional unit supported a more balanced view of AI as both useful and socially consequential. Such a result is compatible with prior scholarship arguing that AI ethics and social impact should not be treated as optional add-ons in K–12 AI education, but rather as integral parts of what students should learn when engaging with AI systems [15], [16].

The increase in efficacy of AI may be understood in light of repeated mastery experiences across the instructional sequence. Students assembled the robot arm, tested basic control, trained image-classification models, revised the models through data improvement, and finally applied the models to a functioning physical system. These repeated cycles of building, testing, revising, and applying likely strengthened students' confidence that they could understand and work with AI-related tasks. This interpretation is also consistent with PBL literature emphasizing that students' sense of competence develops when they engage in authentic, structured inquiry supported by appropriate scaffolding and aligned assessment [17]–[20].

The present findings also support the educational value of combining AI learning with robotics and physical computing. Educational robotics research has shown that robotics can positively influence motivation and achievement, but that its effectiveness depends heavily on task design, teacher support, and the role assigned to the robot in the learning process [21], [22]. In the present study, the robot was not used merely as a motivational device or a programmable artifact. Rather, it served as a physical interface through which students could externalize and evaluate AI decisions. This distinction appears important because the educational value of the unit lay not only in building the robot arm, but in making the consequences of AI outputs physically observable and therefore open to critique, reflection, and redesign [21], [22].

A particularly important contribution of the present study lies in its deliberate alignment of technical design with school reality. The choice of a 2-DOF structure was not merely a simplification for engineering convenience; rather, it was an instructional and implementation decision intended to reduce assembly difficulty, lower the burden of control logic, and make embodied AI activities more accessible to novice learners in ordinary middle-school classrooms. In addition, the USB single-cable design and power-optimization strategy allowed the kit to operate without a separate external power supply, which is especially meaningful in school settings where older PCs, limited outlets, and constrained preparation time often make robotics activities difficult to implement at scale. From this perspective, the novelty of the present study lies not only in developing a low-cost robotic arm kit, but in designing a school-scalable embodied AI learning platform that is technically feasible under resource-constrained classroom conditions [23]–[25].

This practical contribution becomes clearer when viewed against the broader literature on physical computing and low-cost classroom technologies. Previous studies have suggested that microcontroller-based instruction and Arduino-centered learning can support programming, computational thinking, and STEM learning when activities are feasible, accessible, and aligned with classroom realities [23]–[25]. The present study extends this line of work by showing that low-cost physical computing can also serve as a platform for embodied AI learning in middle school technology education. In particular, the 2-DOF configuration and USB single-cable design appear to have lowered the technical burden of implementation while maintaining meaningful opportunities for students to examine model error, uncertainty, and physical misoperation. This suggests that low-cost design is not merely a practical constraint but also a pedagogical strength when it supports classroom manageability, interpretability, and broader dissemination of robotics and AI education in under-resourced school settings [23]–[25].

Another important contribution of the study lies in the integration of AI learning with responsible and governance-oriented reasoning. The qualitative results showed that students did not simply focus on successful classification or efficient automation; they also reflected on malfunction, surveillance, misuse, maintenance burden, emergency-stop systems, safety zones, inspection procedures, and data/model updating. This indicates that students were beginning to understand that safe AI use depends not only on hardware protection, but also on ongoing data and model management. Such findings are aligned with recent calls to embed socio-ethical reasoning directly into AI learning experiences and to support students in thinking about accountability, fairness, risk, and governance alongside technical functionality [5], [15], [16].

The present findings are also relevant to the Korean curricular context. As noted earlier, middle school AI education in Korea has often been implemented mainly within informatics and computing subjects, while opportunities for extended, hands-on AI application in technology education remain limited [28]. Even when physical computing is used, activities are often centered on rule-based programming rather than on the full AI pipeline. The present study therefore contributes to the field by showing that middle school technology education can provide an effective setting for embodied AI learning in which students engage with both engineering constraints and AI-system behavior. In this respect, the study complements earlier Korean work on robot-arm fabrication and control [29] and AI-based invention education [30] by offering a more integrated classroom model that links model outputs, physical actions, and responsible design considerations within one instructional sequence [28]–[30].

The qualitative findings further strengthen the interpretation of the quantitative results. Students identified four overarching themes—perceived applicability, expected benefits, concerns and risks, and rules for safe and responsible use—and these themes help explain why the experimental group showed stronger gains than the comparison group. Students in the AI-integrated condition were not only assembling and controlling a robot arm, but also evaluating model behavior, anticipating failure, and proposing safeguards. This qualitative pattern supports the view that the unit promoted AI literacy beyond technical familiarity toward evaluative and responsibility-oriented reasoning. It also reinforces the claim that the educational value of the intervention lay in embedding the full AI pipeline within a constrained but meaningful physical design task rather than in robotics activity alone [6], [7], [15], [16].

Despite these strengths, several limitations should be acknowledged. First, the study was conducted in one school using four intact classes, and students were nested within class units. Although baseline equivalence was established and ANCOVA was used to control for pretest differences, the class-based structure means that independence assumptions may not have been fully satisfied. Second, the study examined short-term effects within an eight-lesson unit, and therefore the long-term durability of the observed gains remains unknown. Third, the quantitative outcomes relied primarily on self-report measures, and the dimensional structure of the instrument was not revalidated in the present sample. Fourth, although the qualitative analysis was strengthened by a clear analytic framework, future research should enhance trustworthiness further through double-coding, triangulation, or peer debriefing [10], [12], [31].

Overall, the present study has both theoretical and practical significance. Theoretically, it supports the view that embodied interaction with a physical system can deepen AI literacy by making the consequences of model performance visible, discussable, and revisable. Practically, it offers a feasible model for middle school technology education that combines low-cost robotics, project-based design activity, and responsible AI reflection under ordinary school conditions. Rather than claiming that robotics alone guarantees better AI learning, this study suggests that meaningful outcomes emerge when robotics is used to externalize the AI pipeline and when students are supported in examining both the possibilities and the limitations of AI systems through structured design, evaluation, and reflection [17], [18], [21], [22].

5. CONCLUSION

This study developed a low-cost, classroom-ready 2-DOF robotic arm kit and an AI-integrated instructional unit for middle school technology education, and examined both its classroom feasibility and its educational effects. The findings suggest that the developed kit and accompanying materials were feasible and appropriate for implementation in an ordinary school setting. Expert review and pilot implementation observations indicated that the instructional package was manageable for middle school learners and suitable for supporting hands-on AI learning under typical classroom constraints.

Participation in the unit improved students' AI-related outcomes more effectively than conventional technology education. After controlling for pretest scores, the experimental group significantly outperformed the comparison group in value of AI, efficacy of AI, AI literacy, and total score. In addition, qualitative findings showed that students not only recognized the applicability and benefits of AI, but also reflected critically on risks such as malfunction, misuse, surveillance, job displacement, and the need for layered safety governance. These results suggest that AI learning becomes more educationally powerful when it moves beyond isolated coding or tool use and instead engages learners in the full cycle of data collection, model training, evaluation, revision, and real-world application.

The study has both theoretical and practical significance. Theoretically, it shows how embodied interaction with a physical system can deepen AI literacy by making the consequences of model performance visible and discussable. Practically, it offers a school-implementable instructional model that combines low-cost robotics, design-based activity, and responsible AI reflection in a way that is accessible to middle school classrooms. In particular, the constrained 2-DOF system appears to have provided an effective balance between simplicity, hands-on engagement, and meaningful exposure to AI-system limitations. Additionally, future research should examine the reproducibility of the intervention across more diverse schools and grade levels, test longer-term effects through delayed posttests, and employ performance-based rubrics and multilevel analytical approaches to provide stronger evidence of learning and transfer.

FUNDING INFORMATION

Authors state no funding involved.

AUTHOR CONTRIBUTIONS STATEMENT

This journal uses the Contributor Roles Taxonomy (CRediT) to recognize individual author contributions, reduce authorship disputes, and facilitate collaboration.

Name of Author	C	M	So	Va	Fo	I	R	D	O	E	Vi	Su	P	Fu
Dasol Kim	✓	✓	✓	✓	✓	✓		✓	✓	✓	✓	✓	✓	
Sooin Kim		✓	✓	✓		✓	✓			✓	✓			

C : **C**onceptualization	I : **I**nvestigation	Vi : **Vi**sualization
M : **M**ethodology	R : **R**esources	Su : **Su**pervision
So : **So**ftware	D : **D**ata Curation	P : **P**roject administration
Va : **Va**lidation	O : Writing - **O**riginal Draft	Fu : **Fu**nding acquisition
Fo : **Fo**rmal analysis	E : Writing - Review & **E**diting	

CONFLICT OF INTEREST STATEMENT

Authors state no conflict of interest.

INFORMED CONSENT

We have obtained informed consent from all individuals included in this study.

ETHICAL APPROVAL

The human-subject research was conducted in compliance with relevant national regulations and institutional policies in accordance with the tenets of the Declaration of Helsinki and was approved by the head of the authors' institution (the school principal).

DATA AVAILABILITY

The data that support the findings of this study are available from the corresponding author, [DK], upon reasonable request.

REFERENCES

[1] UNESCO, *Beijing Consensus on Artificial Intelligence and Education*. Paris, France: UNESCO, 2019. [Online]. Available: https://unesdoc.unesco.org/ark:/48223/pf0000368303.

[2] F. Miao, W. Holmes, R. Huang, and H. Zhang, *AI and Education: Guidance for Policy-Makers*. Paris, France: UNESCO, 2021, doi: 10.54675/PCSP7350.

[3] Organisation for Economic Co-operation and Development (OECD), Recommendation of the Council on Artificial Intelligence, OECD/LEGAL/0449, 2019. [Online]. Available: https://legalinstruments.oecd.org/en/instruments/oecd-legal-0449

[4] European Commission High-Level Expert Group on AI, *Ethics Guidelines for Trustworthy AI*. Brussels, Belgium: European Commission, 2019, doi: 10.2759/346720.

[5] National Institute of Standards and Technology (NIST), *Artificial Intelligence Risk Management Framework (AI RMF 1.0)*. NIST AI 100-1, Jan. 2023, doi: 10.6028/NIST.AI.100-1.

[6] D. Long and B. Magerko, "What is AI literacy? Competencies and design considerations," in *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems (CHI'20)*, Honolulu, USA, 2020, pp. 1–16, doi: 10.1145/3313831.3376727.

[7] D. T. K. Ng, J. K. L. Leung, S. K. W. Chu, and M. S. Qiao, "Conceptualizing AI literacy: An exploratory review," *Computers and Education: Artificial Intelligence*, vol. 2, Art. no. 100041, 2021, doi: 10.1016/j.caeai.2021.100041.

[8] D. Touretzky, C. Gardner-McCune, F. Martin, and D. Seehorn, "Envisioning AI for K–12: What should every child know about AI?" in *Proceedings of the Thirty-Third AAAI Conference on Artificial Intelligence and Thirty-First Innovative Applications of Artificial Intelligence Conference and Ninth AAAI Symposium on Educational Advances in Artificial Intelligence*, vol. 33, no. 1, pp. 9795–9799, 2019, doi: 10.1609/aaai.v33i01.33019795.

[9] D. Touretzky, C. Gardner-McCune, and D. Seehorn, "Machine Learning and the Five Big Ideas in AI," *International Journal of Artificial Intelligence in Education*, vol. 33, pp. 233–266, 2022, doi: 10.1007/s40593-022-00314-1.

[10] R. Huang, Y. Yin, N. Zhou, and F. Lang, "Artificial intelligence in K-12 education: An umbrella review," *Computers and Education: Artificial Intelligence*, vol. 10, Art. no. 100519, 2026, doi: 10.1016/j.caeai.2025.100519.

[11] J. Su, Y. Zhong, and D. T. K. Ng, "A meta-review of literature on educational approaches for teaching AI at the K-12 levels in the Asia-Pacific region," *Computers and Education: Artificial Intelligence*, vol. 3, 2022, doi: 10.1016/j.caeai.2022.100065.

[12] D. Wu, M. Chen, X. Chen, and X. Liu, "Analyzing K-12 AI education: A large language model study of classroom instruction on learning theories, pedagogy, tools, and AI literacy," *Computers and Education: Artificial Intelligence*, vol. 7, p. 100295, Dec. 2024, doi: 10.1016/j.caeai.2024.100295.

[13] S.-W. Kim and Y. Lee, "The Artificial Intelligence Literacy Scale for Middle School Students," *Journal of the Korea Society of Computer and Information*, vol. 27, no. 3, pp. 225–238, 2022, doi: 10.9708/jksci.2022.27.03.225.

[14] A. Carolus, J. Binder, and F. Muench, "Measuring AI literacy skills (MAILS): Development and validation of a scale," *Computers & Education*, vol. 206, p. 104903, 2024, doi: 10.1016/j.compedu.2023.104903.

[15] S. Akgun and C. Greenhow, "Artificial intelligence in education: Addressing ethical challenges in K-12 settings," *AI and Ethics*, vol. 2, no. 3, pp. 431–440, 2022, doi: 10.1007/s43681-021-00096-7.

[16] R. Williams *et al.*, "AI + Ethics Curricula for Middle School Youth: Lessons Learned from Three Project-Based Curricula," *International Journal of Artificial Intelligence in Education*, vol. 33, pp. 325–383, 2023, doi: 10.1007/s40593-022-00298-y.

[17] D. Kokotsaki, V. Menzies, and A. Wiggins, "Project-based learning: A review of the literature," *Improving Schools*, vol. 19, no. 3, pp. 267–277, 2016, doi: 10.1177/1365480216659733.

[18] S. Lavado-Anguera, P.-J. Velasco-Quintana, and M.-J. Terrón-López, "Project-Based Learning (PBL) as an Experiential Pedagogical Methodology in Engineering Education: A Review of the Literature," *Education Sciences*, vol. 14, no. 6, p. 617, 2024, doi: 10.3390/educsci14060617.

[19] C.-Y. Wang, B.-L. Gao, and S.-J. Chen, "The effects of metacognitive scaffolding of project-based learning environments on students' metacognitive ability and computational thinking," *Education and Information Technologies*, vol. 29, pp. 5485–5508, 2024, doi: 10.1007/s10639-023-12022-x.

[20] H. Thomann and V. Deutscher, "Scaffolding through prompts in digital learning: A systematic review and meta-analysis of effectiveness on learning achievement," *Educational Research Review*, vol. 47, 2025, doi: 10.1016/j.edurev.2025.100686.

[21] F. Ouyang and W. Xu, "The effects of educational robotics in STEM education: a multilevel meta-analysis," *International Journal of STEM Education*, vol. 11, p. 7, 2024, doi: 10.1186/s40594-024-00469-4.

[22] W. Xu and F. Ouyang, "Robotic roles in education: A systematic review based on a proposed framework of the learner-robot relationships," *Educational Research Review*, vol. 47, p. 100685, 2025, doi: 10.1016/j.edurev.2025.100685.

[23] P. A. García-Tudela and J.-A. Marín-Marín, "Use of Arduino in Primary Education: A Systematic Review," *Education Sciences*, vol. 13, no. 2, p. 134, 2023, doi: 10.3390/educsci13020134.

[24] J.-A. Marín-Marín, P. A. García-Tudela, and P. Duo-Terrón, "Computational thinking and programming with Arduino in education: A systematic review for secondary education," *Heliyon*, vol. 10, no. 8, 2024, doi: 10.1016/j.heliyon.2024.e29177.

[25] E. Lee, "Developing a low-cost microcontroller–based model for teaching and learning," *European Journal of Educational Research*, vol. 9, no. 3, pp. 921–934, 2020, doi: 10.12973/eu-jer.9.3.921.

[26] H. Montiel and M. G. Gomez-Zermeño, "Educational Challenges for Computational Thinking in K–12 Education: A Systematic Literature Review of 'Scratch' as an Innovative Programming Tool," *Computers*, vol. 10, no. 6, p. 69, 2021, doi: 10.3390/computers10060069.

[27] M. Carney *et al.*, "Teachable Machine: Approachable web-based tool for exploring machine learning classification," in *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems (CHI EA'20)*, 2020, pp. 1–8, doi: 10.1145/3334480.3382839.

[28] M.-y. Kim and C.-h. Lee, "Exploratory Data Analysis on the Connection of the Software Curriculum between the Primary and Secondary Curriculums and the Higher Curriculums," *Journal of Practical Engineering Education*, vol. 16, no. 3, pp. 283–290, 2024, doi: 10.14702/JPEE.2024.283.

[29] E.-Y. Chang and H.-J. Yu, "A study on the Development of Fusion Education Attempting to Utilize 3D Printing for the Fabrication and Control of Robot Arms," *Journal of Practical Engineering Education*, vol. 16, no. 2, pp. 121–128, 2024, doi: 10.14702/JPEE.2024.121.

[30] S. J. Park, "Development and Effectiveness of Invention Gifted Education Program Using AI," *Korean Association for Learner-Centered Curriculum and Instruction*, vol. 23, no. 8, pp. 81–95, 2023, doi: 10.22251/jlcci.2023.23.8.81.

[31] V. Braun and V. Clarke, "Using thematic analysis in psychology," *Qualitative Research in Psychology*, vol. 3, no. 2, pp. 77–101, 2006, doi: 10.1191/1478088706qp063oa.

BIOGRAPHIES OF AUTHORS

Dasol Kim earned a Ph.D. in Education with a major in Industrial Technology Education from Chungnam National University. He has worked as an educator and researcher for 12 years at institutions including Daejeon Daecheong Middle School, Daejeon Saemirae Middle School, the Daejeon Institute of Education Policy, and Chungnam Middle School. He currently serves as a lecturer at the College of Creativity and Convergence at Chungnam National University and as a lecturer in Digital Education at the Teacher Education Department of Mokwon University. His main research interests include technology education, invention education, maker education, AI-integrated education, robotics education, digital education, and career education. He can be contacted at email: losad721@naver.com.

Sooin Kim earned an M.Ed. in Creativity, Invention, and Convergent Science Education from Seoul National University of Education. He currently serves as the CEO of Creativity & Education Co., Ltd. and as a board member of the Korean Society for Creativity. He has been dedicated to advancing creativity education, robotics education, invention education, and maker education through the development and consulting of creative educational content, training creative education instructors, operating experiential learning programs, and distributing educational products. He can be contacted at email: todaksen@naver.com.