2026 ASEE ANNUAL CONFERENCE June 21–24, 2026 | Charlotte Convention Center, NC
Host Sponsor UNC Charlotte
Paper ID #53371

AI core competencies evaluation for elementary and secondary school students: model, framework, and application

Prof. Xiwei Liu, State Key Laboratory of Multimodal Artificial Intelligence Systems, Institute of Automation, Chinese Academy o

Xiwei Liu is an associate professor at State Key Laboratory of Multimodal Artificial Intelligence Systems, Institute of Automation, Chinese Academy of Sciences. He received his doctoral degree from Nara Institute of Science and Technology (NAIST) in Japan. His current research interests mainly focus on AI in education, computer vision, educational assessment, and STEM education.

DAN ZHOU, Future Gene (Beijing) Artificial Intelligence Research Institute Co., Ltd.

An expert in artificial intelligence education, he has long focused on the assessment of adolescents' AI literacy and exploration of its influencing factors, with extensive practical experience in the AI industry. He holds a master's degree from the Department of Intelligent Science, Peking University, and has conducted doctoral research at the School of Education, Tsinghua University.

Miss YIHAN ZHOU, Capital Normal University, Beijing, China

Yihan Zhou is a Master's student in the Faculty of Education at Capital Normal University and concurrently serves as a Research Assistant at the Institute of Automation, Chinese Academy of Sciences. Her research focuses on exploring new pathways for empowering education with intelligent technologies within the framework of curriculum and instruction theory.

Chao Gong

AI core competencies evaluation for elementary and secondary school students: model, framework, and application

Abstract

There are rapid progresses and extensive application of artificial intelligence in almost all kinds of industries in recent decades. In 2017, Next Generation Artificial Intelligence Development Plan was Issued by China's State Council and various forms of activities to popularize AI were encouraged and supported in the plan. Especially, AI education programs were proposed including AI-related courses in elementary and secondary schools. Plenty of devices, robots, integrated development kits with series of curriculums and books were developed and supplied by companies to various training classes inside and outside schools.

However, it's an urgent need to build a standard curriculum and evaluation methodology for the AI education of elementary and secondary school students. Based on the program of "Definition and Selection of Competencies: Theoretical and Conceptual Foundations" (DeSeCo) proposed by OECD, the core competencies and values for Chinese students ’ development proposed by Beijing Normal University, and curriculum standard of information science and technology proposed by Ministry of Education of China, an new evaluation model for AI core competencies was built by us with four dimensions: 1. AI Disciplinary Literacy and Skills; 2. Higher-Order Thinking in the AI Era; 3. Human-AI Hybrid Collaborative Innovation; 4. AI Social Responsibility.

Furthermore, a practical framework was proposed base on the evaluation model of AI education. There were mainly two key characters of the framework: a). 1 + X, which means the evaluation should be done from one common part, AI fundamentals and several different subjects, such as scratch, python programming, micro:bit or Arduino based application etc. b). step-by-step, which means every evaluation subject includes several steps, 1st-8th levels generally, and the evaluation was also divided into two phases, popularization and deep application. In conclusion, the evaluation framework had many different subjects with multiple dimensions.

The AI core competencies evaluation (AICE) program has been carried out once a quarter since November 2021. More than 75,000 primary and secondary school students have participated in the program, and many of these students have passed six levels of the subject. Two study cases are illustrated in this paper. The first case focuses on the empirical application of the AICE project, elaborating on the implementation and implications of the AICE evaluation. The second case presents an analysis of test data regarding the changes in 20 high school students ’ performance before and after participating in a smart car course within a semester, so as to assess the improvement of their AI core competencies.

Keywords: information literacy, Multilevel Program Assessment, Elementary and middle school, STEM, case study

1. Introduction

Artificial Intelligence (AI) is gradually evolving into the "underlying technical logic and basic cognitive grammar" that reshapes social production and lifestyle, as well as reconstructs the educational ecosystem[1]. This transformation requires the education sector to redefine and construct a core competency system adapted to the intelligent era. As one of the core digital skills in the 21st century, AI literacy has become an indispensable core competency for future citizens to adapt to social development and achieve lifelong learning, alongside traditional foundational literacies such as reading, writing, and mathematics[2]. From the perspective of the evolution of educational interaction forms, AI will reshape the relational dimensions in the educational field in multiple roles: as an auxiliary teaching partner to reconstruct teacher-student interaction models, as a personalized learning carrier to innovate students' self-cognition and development paths, and as a collaborative medium to optimize peer interaction mechanisms. This human-machine collaborative and symbiotic educational ecosystem places higher demands on learners' critical thinking and technical ethics awareness, making it urgent and important to proactively explore how to cultivate youth competencies to adapt to and harness AI[3].

Given the extensive penetration of AI technology in various industries and its crucial significance for future social development, UNESCO clearly stated in its publication AI and Education: Guidance for Policy-Makers[4]: "All citizens need to possess a certain level of AI literacy to adapt to an increasingly AI-driven world," elevating AI literacy to the height of a basic right of global digital citizens. The AI Competency Framework for Students (AICFS) released in 2024[5] further designates AI literacy as an essential competency for both students and teachers, and for the first time constructs a matrix-based competency structure of "four horizontal elements + three vertical levels," providing a unified reference for global education practices across different academic stages. China has established a policy system for AI education that progresses from pilot exploration to popularization: the National New Generation AI Development Plan in 2017 first proposed "setting up AI-related courses in primary and secondary schools," marking the initial development stage of AI education; the General High School Information Technology Curriculum Standards issued in the same year included "Introduction to Artificial Intelligence" as an optional compulsory module; the Compulsory Education Information Technology Curriculum Standards in 2022 centered information technology teaching around the main line of AI; and the Notice on Strengthening AI Education in Primary and Secondary Schools in 2024 further clarified the goal of "basically popularizing AI education in primary and secondary schools by 2030" and proposed incorporating AI literacy into students' comprehensive competency evaluation. China has thus become one of the few countries in the world that promotes AI education in primary and secondary schools through national strategic coordination[6].

Against this backdrop, this study focuses on "the construction of an AI competency model, design of an assessment framework, and empirical application for Chinese youth," aiming to address the current problems of fragmented frameworks, lack of assessment tools, and insufficient practical implementation in existing research. It provides theoretical basis and practical reference for the optimization of AI education policies in China and globally, as well as the improvement of youth competency levels. The structure of the article is as follows: Section 2 reviews domestic and international research on AI literacy and evaluation; Section 3 elaborates on the AI core competencies evaluation model; Section 4 presents the "1+X" stepped assessment framework; Section 5 analyzes the application effects through empirical research on the AICE project; Section 6 discusses the research results, limitations, and future research directions.

2. Literature Review

2.1 Definition of Core Competencies

The systematic exploration of literacy and core competencies internationally began in the late 20th century: in 1997, OECD launched the "Definition and Selection of Key Competencies (DeSeCo)" project, defining competencies as "key abilities and essential qualities that individuals need to adapt to society, engage in lifelong learning, and meet future challenges," emphasizing contextualization and comprehensiveness. In 2018, UNESCO further expanded the connotation of literacy, proposing that it is "a comprehensive ability of individuals to identify and understand information, create content, and achieve effective communication and computation with the help of context-related materials," highlighting the practical attributes in a technological environment. In China, the Ministry of Education (MOE) released the Overall Framework of Chinese Students' Core Competencies in 2016, centering on "the all-round development of individuals" and dividing competencies into three major domains: Cultural Foundation (humanistic heritage, scientific spirit), Independent Development (learning to learn, healthy living), and Social Participation (responsibility and commitment, practical innovation). This framework not only focuses on individual growth but also responds to social needs, providing an underlying logic for the construction of subsequent disciplinary competencies and cross-disciplinary competencies (such as AI literacy).

2.2 Research Progress on AI Literacy Frameworks

International research on AI literacy has shown an evolutionary trend from single skills to comprehensive competencies: Long and Magerko (2020) first proposed a three-dimensional framework from the perspective of human-computer interaction — "critically evaluating AI technologies, effectively communicating and collaborating with AI, and using AI as a tool"—and refined 17 specific abilities, laying the theoretical foundation for AI literacy research[7]; UNESCO's (2022) K-12 AI Curriculum: A Blueprint for Government-Recognized AI Education divided AI literacy into three major dimensions: "knowledge, skills, and affective values," becoming the first global AI literacy framework[8]; the 2024 AICFS further innovated a structure of "four horizontal elements (knowledge, skills, attitudes, ethics) + three vertical levels (foundation, advancement, innovation)," emphasizing spiral curriculum sequences and competency transferability[5]; OECD and the European Commission (2025) decomposed literacy into four major domains from a practical orientation: "Engaging with AI (critically examining outputs), Creating with AI (collaborative innovation), Managing AI (task adaptation and ethical regulation), and Designing AI (exploring technical principles)," highlighting the orientation of green transformation and sustainable development[6].

Research on AI literacy frameworks in China has mainly formed two paths: (1) The disciplinary competency transfer path, influenced by the formulation of existing disciplinary competency dimensions, elaborates on the conceptual connotation of AI literacy from the perspectives of learners' knowledge, skills, thinking, and ethical responsibilities. For example, the Chinese Association of Automation (CAA) divided youth AI core competencies into "disciplinary literacy and skills, higher-order thinking in the intelligent era, human-machine hybrid collaborative innovation, and intelligent social responsibility" in 2021, breaking through the previous limitation of focusing on programming as the core and computational thinking as the sole connotation, and strengthening the cultivation of higher-order thinking and collaborative innovation[9]; the MOE's General Guidelines for AI General Education in Primary and Secondary Schools (2025) proposed a four-in-one literacy of "scientific and technological innovation thinking, critical thinking, human-machine collaboration ability, and social responsibility awareness," further connecting with the national core competency framework[10]. (2) The literacy definition derivation path: the Framework for AI Technology and Engineering Literacy in Primary and Secondary Schools developed by the Central Institute of Educational Technology in 2022 referenced OECD's ideas and expanded from four dimensions: "AI and humans, AI and society, AI technology, and AI system design." The content description is relatively abstract, but there is a corresponding conversion relationship with the disciplinary competency transfer path. For example, "AI and tools/technology" reflects how humans reshape their relationship with new tools in an intelligent society, corresponding to dimensions such as "AI knowledge and AI ability"; "AI and self" emphasizes individuals' autonomy, self-cognition, and adaptability in an intelligent environment, corresponding to dimensions such as "AI awareness, AI thinking, and AI emotion"; "AI and humans/society" emphasizes the complex relationship and ethics between AI and humans, as well as the profound impact on social development, corresponding to dimensions such as "AI ethics."

In summary, there is still no unified definition of AI literacy in the academic community, with fragmented framework definitions presenting a "thousand people, thousand faces" phenomenon. Most studies remain at the level of theoretical indicator analysis and exploration, or are developed from an "engineer's perspective"[11], lacking sufficient practical experience support in primary and secondary school applications, which may lead to difficulties in implementation. The youth AI literacy referred to in this study refers to the comprehensive ability of youth to identify, understand, use, create, and critically evaluate various AI technologies in daily study, life, and social interactions, while following correct values and ethical norms, assuming civic social responsibilities, and possessing lifelong learning capabilities. It is a comprehensive quality integrating knowledge, skills, thinking, practice, and responsibility to adapt to the intelligent era.

3. AI Competency Evaluation Model

The youth AI core competency framework proposed in this study is divided into four dimensions: AI Disciplinary Literacy and Skills, Higher-Order Thinking in the AI Era, Human-AI Hybrid Collaborative Innovation, and AI Social Responsibility. This framework has a certain corresponding relationship with the three dimensions ("Cultural Foundation, Independent Development, Social Participation") in the Overall Framework of Chinese Students' Core Competencies issued by the MOE of China, as shown in Figure 1.

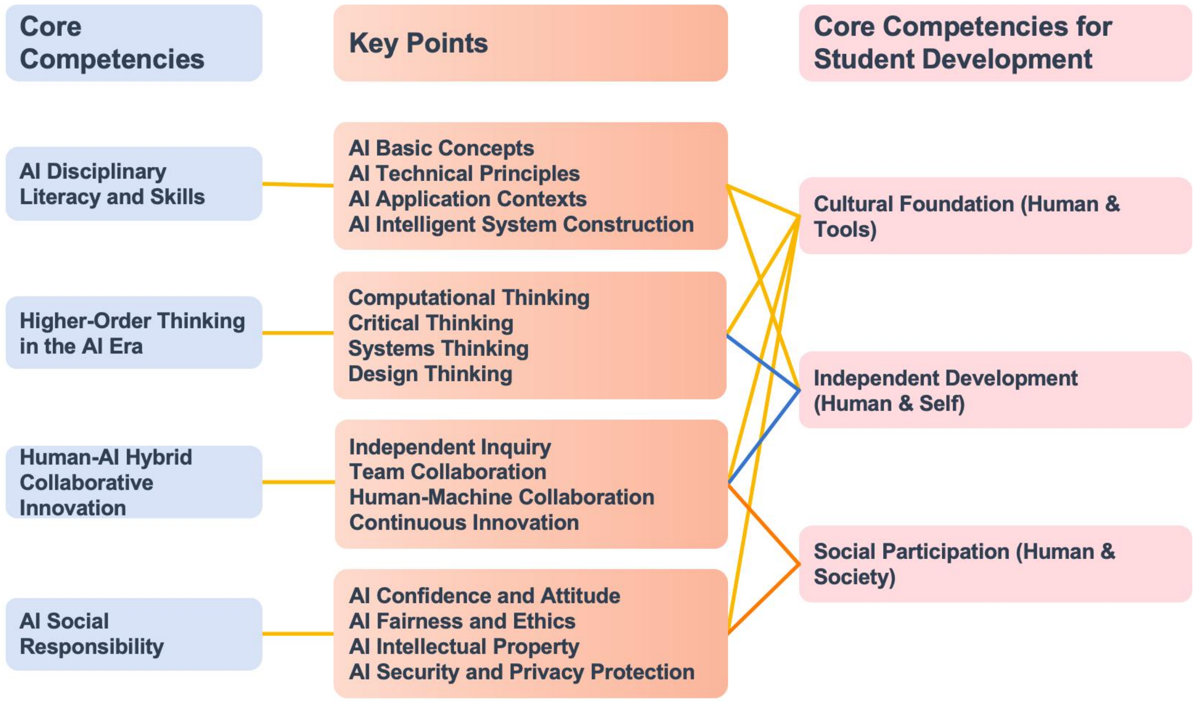

Figure 1 Artificial Intelligence Core Competency Model for Youth

3.1 Disciplinary Literacy and Skills

As the foundation of AI literacy, this dimension focuses on the mastery of AI knowledge and technical application capabilities, forming a progressive logic across academic stages:

At the primary school stage, the evaluation focuses on AI technology identification and basic concept cognition, such as recognizing AI products in daily life (smart speakers, facial recognition) and understanding AI application scenarios (education, medical care), requiring mastery of basic terms such as "data, algorithms, and models";

At the junior high school stage, the evaluation focuses on initial technical operation and problem-solving abilities, such as using graphical programming (Scratch) to implement simple interactive functions, completing sensor data collection through Arduino, and understanding the basic process of machine learning;

At the senior high school stage, the evaluation focuses on technical integration and interdisciplinary application, such as processing data with Python, optimizing algorithms with C++, and designing AI-assisted inquiry projects (e.g., intelligent temperature control systems) in combination with mathematics and physics knowledge.

3.2 Higher-Order Thinking in the Intelligent Era

As the core of AI literacy, this dimension focuses on the cultivation of computational thinking, systems thinking, critical thinking, and design thinking. Research estimates that 65% of current primary school students will engage in occupations that do not exist today.

At the primary school stage, the key evaluation content is problem decomposition and logical reasoning abilities, such as decomposing the "intelligent garbage classification" task into steps of "image recognition-category matching-prompt feedback";

At the junior high school stage, the focus is on system analysis and critical examination abilities, such as analyzing the information cocoon effect of AI recommendation algorithms and proposing optimization suggestions;

At the senior high school stage, the emphasis is on innovative design and complex problem-solving abilities, such as designing an intelligent campus energy consumption management system that comprehensively considers technical feasibility and cost-effectiveness.

3.3 Human-Machine Hybrid Collaborative Innovation

As an extension of AI literacy in practice, this dimension focuses on human-machine collaboration and team innovation capabilities. The landmark event of AI defeating top human players in Go indicates that the living and development environment of humans has expanded from natural and social environments to machine and virtual environments.

At the primary school stage, the evaluation focuses on human-machine collaboration awareness, such as assisting creative expression through AI painting tools and understanding the relationship between human instructions and machine outputs;

At the junior high school stage, the focus is on team collaboration and task division abilities, such as completing the development of an intelligent campus navigation robot through group cooperation and clarifying role assignments;

At the senior high school stage, the emphasis is on cross-team collaboration and the ability to implement innovative outcomes, such as designing a rural smart agriculture monitoring system by integrating multidisciplinary teams to meet practical needs.

3.4 Intelligent Social Responsibility

As the ethical bottom line of AI literacy, this dimension focuses on AI ethical norms, social emotions, and responsibility awareness. Rational application, appropriate control, and standardized management can effectively mitigate potential risks.

At the primary school stage, the evaluation focuses on initial cognition of AI ethics, such as recognizing the risks of AI face swapping and establishing the awareness of not spreading false AI content;

At the junior high school stage, the focus is on ethical judgment and norm-abiding abilities, such as respecting intellectual property rights in AI-assisted writing;

At the senior high school stage, the emphasis is on ethical construction and social responsibility, such as investigating cases of AI algorithmic discrimination and proposing data fairness optimization solutions.

4. "1+X" Stepped Assessment Framework

Based on the aforementioned competency model, this study constructs a "1+X" composite assessment system: "1" refers to the comprehensive theoretical competency assessment of AI, focusing on the standardized and objective evaluation of youth basic cognition of AI and information literacy, as well as professional module knowledge; "X" corresponds to the hierarchical assessment of AI application and practical abilities, targeting propositions based on the competency dimensions of different practical modules. It covers software programming modules such as graphical programming, Python programming, and C++ programming, as well as hardware-software collaborative programming modules such as Arduino, block robots, and drones. Each practical module assesses youth ability levels in a multi-dimensional and comprehensive manner around the four core competency dimensions: knowledge and skills, higher-order thinking, human-machine hybrid collaborative innovation, and intelligent social responsibility.

The Artificial Intelligence Competencies Evaluation (AICE) for youth has a standardized operation model of national unified proposition, unified assessment management, and unified certificate issuance, and establishes a hierarchical assessment structure of "Foundation Popularization-Advancement Improvement":The Foundation Popularization stage takes standardized paper-based/computer-based tests as the core assessment method, focusing on the compliance assessment of basic literacy;The Advancement Improvement stage adopts a comprehensive assessment form of "project submission + defense." The specific process is as follows: youth complete practical works (including physical objects, schemes, or prototypes) and supporting research reports, then participate in on-site defenses. Review experts conduct targeted questioning and interactive communication to focus on assessing the technical implementation quality of the works, the academic standardization of the reports, and the externalized performance of higher-order thinking abilities (see Tables 1 and 2 for specific configurations). This assessment model breaks through the limitations of traditional single paper-based tests, responds to the cultivation logic of core competencies through hierarchical and progressive design, and can achieve a comprehensive and scientific characterization of youth comprehensive AI literacy levels.

Table 1 Description of Projects in the Foundation Popularization Stage of the "1+X" Assessment System

"X" Assessment Projects	Tools/Modules	Competency Levels
A. AI Programming	1. Graphical Programming Language	Level 1 to Level 5
	2. Python Language	Level 1 to Level 8
	3. C++ Language	Level 1 to Level 5
B. AI Perception and Control	1. Logical Thinking	Level 1 to Level 3
	2. Block Robots	Level 1 to Level 4
	3. Arduino	Level 1 to Level 5
	4. Micro:bit	Level 1 to Level 5
C. AI Product Structure	Structural Construction	Level 1 to Level 4

Table 2 Description of Projects in the Advancement Improvement Stage of the "1+X" Assessment System

"X" Assessment Projects	Tools/Modules	Assessment Methods and Outcomes
D. AI Application Fields	1. Intelligent Transportation 2. Smart Home 3. Smart Agriculture 4. Smart Medical Care	Project Demonstration + Defense(Outcomes include physical objects/schemes/prototypes and research reports)
E. AI Algorithm Foundation	1. Big Data 2. Machine Vision 3. Natural Language Processing 4. Machine Learning	Project Demonstration + Defense(Outcomes include physical objects/schemes/prototypes and research reports)

"X" Assessment Projects	Tools/Modules	Assessment Methods and Outcomes
	5. Knowledge Engineering	

5. Empirical Application of the AICE Project

5.1 Analysis on the Effect of Large-scale Continuous Learning under the Framework of "1 + X"

The "1 + X" teaching framework proposed by this research organization has been promoted on a pilot basis since 2021. After more than five years of practice iteration, it has formed a mature paradigm of "standardization of basic evaluation, individualization of advanced testing and gradient of evaluation system", and has been evaluated every academic year.

The framework's AI Literacy assessment system has covered 34 provincial-level administrative regions in China, achieving nationwide coverage. With a total of 85,276 students participating in 2025, the study has formed a super-large-scale dataset reflecting teaching practices across all school-age groups, regions, and multiple practice directions. It provides a solid practical basis for empirical research.

As illustrated in Figure 2 regarding the age distribution of participants, the assessment prioritized core school-age groups in primary and secondary education. Specifically, students in the upper grades of primary school constituted the largest segment. The total proportion of students in the core age range of 8 – 12 years old reached 85.07% (n=72530), with 9-year-olds being the most prevalent subgroup at 19,690 (23.09%). Conversely, the younger group (5 – 7 years old) and the older group (13 – 16 years old) accounted for 4.82% and 10.09%, respectively, while participants over 16 years old represented less than 0.1%. This distribution confirms the framework ’ s precise focus on the core audience of K-12 AI education, establishing a robust and representative foundation for large-scale empirical research.

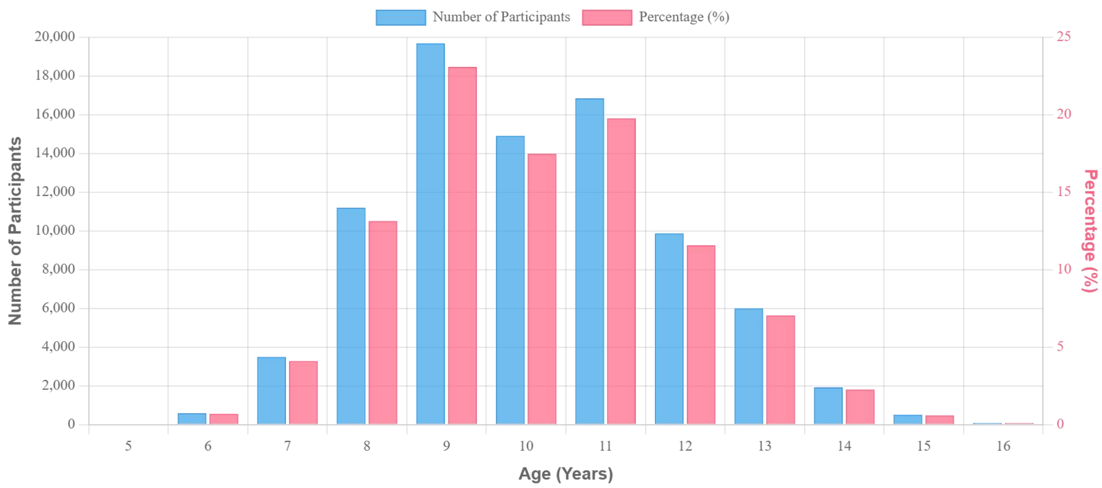

Figure 2 Age distribution of participants in the "1+X" AI education framework assessment

Data sources and sample characteristics

In order to give consideration to the representativeness of the sample and the depth of analysis, this study focuses on the evaluation results of the Python programming direction in the core literacy test in 2025, and takes 2517 students from Jiangsu, Shandong, Zhejiang, Fujian, Anhui and Jiangxi provinces as examples to explain the results, focusing on the analysis of the two major directions of participation frequency and ability level. First of all, the 2517 students who participated in the test can be divided into two types: continuous participation and partial participation: 2239 students who continuously participated in three assessments of Python programming level 1 (basic difficulty) to level 3 (advanced difficulty), and they are also the core samples of the "1 + X" framework effectiveness analysis; Some of the participating students are those who have only participated in one or two Python programming assessments, totaling 278 students. The percentage is shown in Figure 3. Nearly 89% of the students participated in all the gradient assessments from level 1 to level 3, indicating that under the "1 + X" framework, the vast majority of students can continue to participate in the assessment with the teaching progress, and the stability and continuity of vertical participation are excellent; Only about 11% of the students (n=375) participated in 1-2 assessments, of which 2.9% participated in only one assessment and 8.3% participated in two assessments. The low dropout rate reflects that the retention rate and attractiveness of the assessment system are better. The design logic of "foundation + practice" of the framework can effectively maintain students' willingness to participate, and students' participation behavior presents the remarkable characteristics of "high stability and strong sustainability", which fully verifies the attractiveness and adaptability of the framework evaluation system.

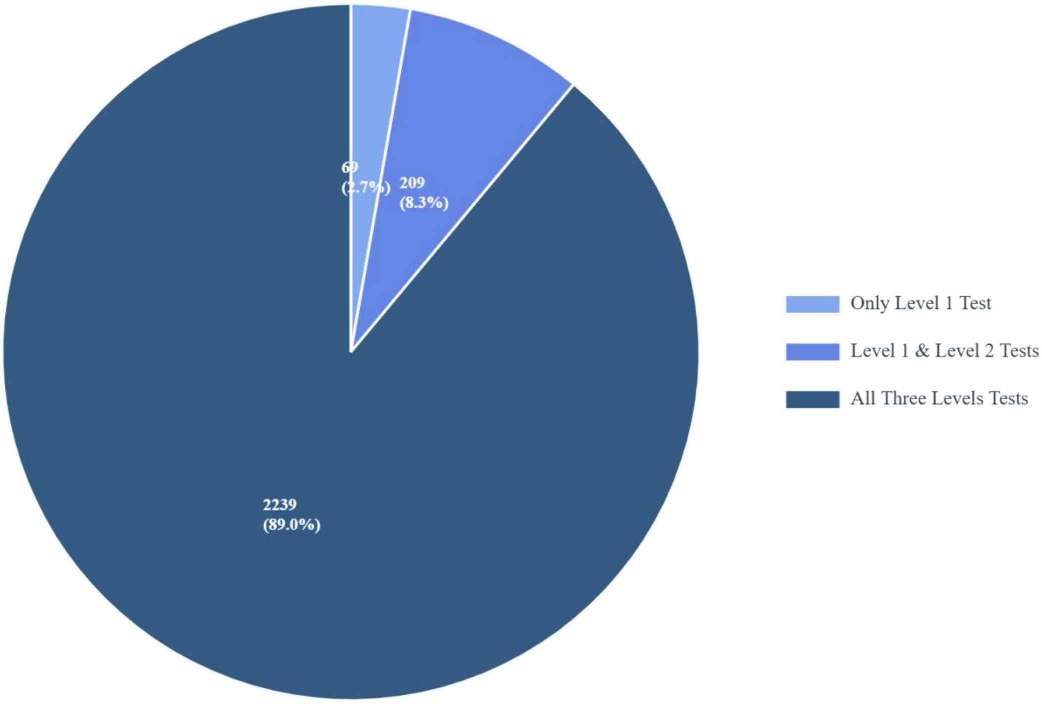

Figure 3 Distribution of Test Participation

Secondly, the grade score distribution of the students who participated in the three assessments is as shown in Figure 4. The number of excellent students in the first level test is 1163, the excellent rate is 51.9%; the number of qualified students is 1075, the qualified rate is 48.0%; the overall passing rate is 99.9%. The second-level assessment focused on program logic and module application ability, and the difficulty was improved compared with the first-level assessment. The number of excellent people was 867, and the excellent rate was 38.7%; the number of qualified people was 1371, the qualified rate was 61.2%, and the overall pass rate was 99.9%. The three-level assessment focuses on comprehensive application and problem solving ability, which is the most difficult assessment among the three levels. In the three-level evaluation, there were 739 excellent people, the excellent rate was 33. 0%; there were 1431 qualified people, the qualified rate was 63. 9%; there were 69 unqualified people, the unqualified rate was 3. 1%; the overall pass rate was 96. 9%.

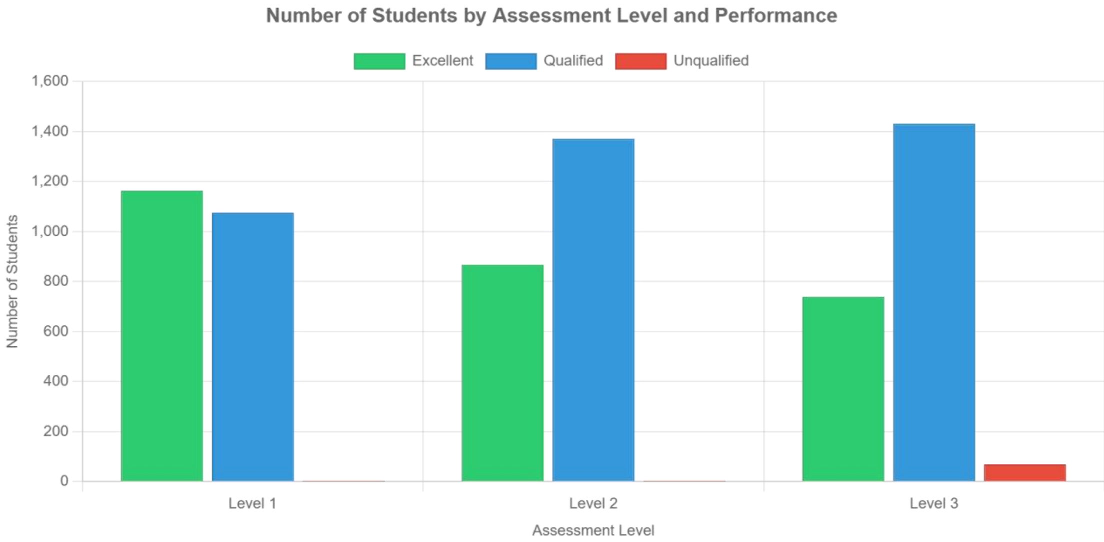

Figure 4 Three-Level Assessment Performance Analysis

From the perspective of trends in the excellent rate (i.e., high-achievement rate), there was a steady decline from Level 1 to Level 3, with a total decrease of 18.9 percentage points. This trend aligns with the objective laws of skill acquisition; as the depth of knowledge points and the complexity of application increase, the proportion of high-performing students naturally diminishes. This outcome validates the rationality of the difficulty gradient designed within the assessment framework. Conversely, the pass rate exhibited a steady upward trend, increasing by a total of 15.9 percentage points. This indicates that the majority of students were able to master the core knowledge points at each level through continuous learning, maintaining a competent level of performance. This reflects the strong overall effectiveness of the teaching design within the "1+X" framework. Furthermore, the pass rate for the Level 3 assessment stabilized at approximately 97%, suggesting a high degree of alignment between the assessment difficulty and the students' actual proficiency levels, without exceeding their capabilities. In summary, the test results demonstrate that the "1+X" framework instructional design centers on competency development. By clarifying stage-specific goal boundaries, optimizing the articulation of knowledge gradients, and reinforcing practical application orientation, the framework constructs a consistent closed-loop system of "Teaching—Learning—Assessment." This approach not only ensures the solid mastery of foundational knowledge but also facilitates the progressive cultivation of higher-order thinking skills. Beyond providing quantitative evidence for the effectiveness of the framework, this assessment practice verifies its practical value in cultivating core AI competencies among youth, thereby offering a replicable and scalable paradigm for optimizing future instructional designs and perfecting assessment systems in similar courses.

5.2 An Empirical Study on the "1+X" Framework Based on Smart Car Project-Based Courses

Methodology

Based on the "1+X" stepped practical framework, this study adopted a quasi-experimental design methodology. Utilizing a smart car project-based learning (PBL) curriculum as the carrier, an experimental intervention was conducted over the course of one semester. The measurement instrument was constructed based on the four core dimensions of AI literacy: AI Disciplinary Literacy and Skills, Higher-Order Thinking in the AI Era, Human-AI Hybrid Collaborative Innovation, and AI Social Responsibility. The number of questions, score weights, and assessment criteria remained consistent between the two tests. However, since some questions involved the cross-examination of 1 to 3 dimensions, the total sum of scores across individual dimensions exceeds the total score of the test paper.

Data collection was conducted in two phases: a pre-test was administered before the curriculum implementation to establish the baseline level of student literacy, while a post-test was conducted after the course concluded to capture changes in literacy following the intervention. During the pre-test statistical analysis, purely rote-memory questions were excluded, establishing an effective full score of 78 points for the paper to ensure the measurement focused on core competencies and literacy. By comparing the pre-test and post-test results, this study analyzed the changes in students' AI literacy to evaluate the impact of the "1+X" teaching framework. The specific metric used was "Experimental Intervention Effect = Post-test Score - Pre-test Score." The course lasted for one semester, yielding a final valid sample of 21 high school students, with the curriculum content designed to systematically examine the improvement of students' AI literacy.

Data analysis employed Paired Samples t-tests, processed using SPSS 27.0. First, the difference between corresponding post-test and pre-test observations was calculated to form a change sequence. The mean of these differences was used to determine the direction of literacy change (a positive value indicating improvement). A significance level of 0.05 was set for hypothesis testing. If the two-tailed significance probability (Sig.) was less than 0.05, the change was considered statistically significant.

Research Results and Analysis

The pre-test results indicated that the average score of students' AI literacy was 46.75 (out of 78), representing a score rate of approximately 59.94%. This suggests that the initial level of students' AI literacy was in the lower-medium range, indicating significant room for improvement. After one semester of intervention using the "1+X" framework, the average post-test score increased to 53.47, with the score rate rising to 68.55%, showing an overall significant positive growth trend. The results of the paired samples t-test, detailed in Table 1, indicated a t-value of -4.474 (df=20) for the total score, with a two-tailed significance probability (Sig.) of < 0.001.This is far below the preset significance level of 0.05, indicating a statistically significant difference in students' overall AI literacy before and after the intervention, thus confirming the significant overall effectiveness of the "1+X" stepped practical framework.

From the perspective of specific assessment results for the four core literacy dimensions as presented in Table 2, different dimensions exhibited distinct improvement characteristics:

1. AI Disciplinary Literacy and Skills:The pre-test mean was 16.5263 (out of 36, score rate 45.91%), and the post-test mean increased to 24.2500 (score rate 67.36%). The t-value was -8.172, and the two-tailed significance probability (Sig.) was 0.001 < 0.05, indicating a statistically significant improvement. These results suggest that through

curriculum intervention, students have systematically mastered the basic concepts, core principles, and technical application logic of AI. They can accurately elaborate on the core processes of machine learning and deep learning, as well as the application boundaries and operational pathways of common AI algorithms in real-world scenarios, reflecting a solid enhancement in foundational literacy.

2. Higher-Order Thinking in the AI Era:The pre-test mean was 33.1429 (out of 52, score rate 63.74%), and the post-test mean reached 36.1667 (score rate 69.55%). The t-value was -2.720, and the two-tailed significance probability (Sig.) was $0.013 < 0.05$, showing a statistically significant improvement. Post-intervention, students demonstrated more proficient use of critical and creative thinking to conduct multi-dimensional analysis and problem-solving regarding complex AI-related issues. For instance, when facing AI ethical dilemmas such as algorithmic bias or data privacy leaks, students could engage in systematic thinking from dimensions such as technical feasibility, moral compliance, and social impact, and propose solutions that are both scientific and practical.

3. AI Social Responsibility:The pre-test mean was 20.5263 (out of 40, score rate 51.32%), and the post-test mean saw a substantial increase to 35.8750 (score rate 89.69%). The t-value was -5.254, and the two-tailed significance probability (Sig.) was $0.001 < 0.05$, representing the most significant improvement margin. This data proves that through course learning, students have developed a profound understanding of the potential positive value and risk implications of AI technology in areas such as employment structures, privacy security, and social equity. They have constructed a moral cognitive system and sense of social responsibility that meets the requirements of the times, and can consciously follow relevant AI ethical guidelines and laws in simulated practical scenarios, laying a solid foundation for technical ethics literacy.

4. Human-AI Hybrid Collaborative Innovation:The pre-test mean was 29.2632 (out of 44, score rate 66.51%), and the post-test mean was 29.7500 (score rate 67.61%). Although showing a slight positive change, the t-value was only -0.329, and the two-tailed significance probability (Sig.) was $0.746 > 0.05$, failing to reach a statistically significant level. Analysis suggests this result is primarily related to the following factors:

   First, the cultivation of human-machine collaborative innovation capabilities requires a solid foundation of practical operational experience; a one-semester intervention duration may be insufficient to meet the cycle requirements for forming such complex capabilities.

   Second, this dimension places high demands on students' interdisciplinary knowledge integration capabilities (such as the fusion of computer programming, mechanical design, data analysis, etc.), whereas students' existing knowledge structures lack sufficient comprehensiveness.

   Third, the cultivation of team collaborative innovation capabilities requires high-frequency, deep-level collaborative practice, and the depth and breadth of relevant task design within the course need to be strengthened.

Table 1 The average score of the students' AI literacy test

		N	Mean	Std. deviation	Std.error mean	t	df	Sig. (2-tailed)
Total Score	pre-test	21	46.7569	6.10641	1.33253	-4.474	20	<0.001
Total Score	post-test	21	53.4702	4.07467	.88917			

Table 2 Assessment Results of the Four Core Literacy Dimensions

		N	Mean	Std. deviation	Std.error mean	t	df	Sig. (2-tailed)
AI Disciplinary Literacy and Skills	pre-test	21	16.5263	3.78640	.82626	-8.172	20	.001
	post-test	21	24.2500	2.88964	.63057			
Higher-Order Thinking in the AI Era	pre-test	21	33.1429	4.84031	1.05624	-2.720	20	.013
	post-test	21	36.1667	4.31663	.94197			
Human-AI Hybrid Collaborative Innovation	pre-test	21	29.2632	4.69938	1.02549	-.329	20	.746
	post-test	21	29.7500	3.62629	.79132			
Human-AI Hybrid Collaborative Innovation	pre-test	21	20.5263	2.88736	.63007	-5.254	20	.001
	post-test	21	35.8750	4.16983	.90993			

Based on the results of the paired samples t-test, it is evident that the two-tailed significance probability (Sig.) for three dimensions—AI Disciplinary Literacy and Skills, Higher-Order Thinking in the AI Era, and AI Social Responsibility—was all less than 0.05. This indicates that after the intervention of the "1+X" AI framework curriculum, students achieved statistically significant improvements in their competencies across these three core literacy dimensions. In contrast, the two-tailed significance probability (Sig.) for the Human-AI Hybrid Collaborative Innovation dimension exceeded 0.05, failing to exhibit a significant improvement effect.

These research findings not only validate the scientificity and effectiveness of the teaching model based on the "1+X" stepped practical framework but also provide a clear target direction for subsequent curriculum optimization. Future studies may focus on enhancing capabilities in Human-AI Hybrid Collaborative Innovation through measures such as extending the duration of practical teaching, promoting interdisciplinary knowledge integration, and designing in-depth collaborative innovation tasks. These efforts aim to promote the comprehensive and balanced development of middle school students' core AI literacy.

6. Conclusions

The study has demonstrated the positive impact of the "1+X" stepped assessment framework on enhancing students' AI competencies, particularly in the areas of general knowledge and skills, higher-order thinking in AI application, and ethics and social responsibility in the AI era. The statistically significant improvements observed in these dimensions underscore the effectiveness of the curriculum intervention in fostering a well-rounded understanding of AI among students.

However, the limited improvement in Human-AI Hybrid Collaborative Innovation highlights the need for further refinement of the teaching model.To address this gap, future curriculum design should consider incorporating longer practical teaching cycles to allow students more time to develop and consolidate their collaborative skills. Additionally, integrating interdisciplinary knowledge more explicitly within the curriculum can help students better understand the complex interactions between different fields in AI applications.

Moreover, designing more in-depth team collaborative innovation tasks can provide students with opportunities to engage in high-frequency, deep-level collaborative practice, thereby enhancing their teamwork and innovation abilities. By implementing these strategies, educators can promote the comprehensive and balanced development of students' core AI competencies, preparing them to meet the challenges and opportunities of the rapidly evolving AI landscape.

References

[1] Cai, Y. C., Zhang, J. B., Yu, C. L., et al. AI Literacy in the Digital Intelligence Era: Connotation, Framework, and Implementation Paths. Journal of Library Science in China, 2024, 50(4):71-84.

[2] Zhang, Y. R., Yang, G., Xu, J. Y., et al. Construction of an AI Literacy Model and Its Implementation Paths. Modern Educational Technology, 2022, 32(3):42-50.

[3] Yoon, J. K., Kim, G., & Smith, A. Playful Design for AI Literacy: Creating Inclusive Learning and Assessment Opportunities. International Journal of Human-Computer Studies, 2025, 199, 103508.

[4] UNESCO. AI and Education: Guidance for Policy-Makers. https://unesdoc.unesco.org/ark:/48223/pf0000378648?locale=en.

[5] UNESCO. AI Competency Framework for Students (AICFS). https://unesdoc.unesco.org/ark:/48223/pf0000391105?posInSet=1&queryId=1ea98642-16e4-4f7d-b432-8536770ad8f5.

[6] Wei, Y. N. Global Experiences and School-Based Paths for Cultivating Students' Digital Literacy. Primary and Secondary School Management, 2024, (5):21-25.

[7] Long, S. D., & Magerko, B.. Computational Thinking and AI Literacy: A Systematic Review of Empirical Studies. Journal of Educational Technology & Society, 2020, 23(4): 1-18.

[8] UNESCO. K-12 AI Curriculum: A Blueprint for Government-Recognized AI Education. Paris: UNESCO, 2022.

[9] Zhang, Y., Luo, X., Guo, Y. X., et al. The Evolution of Aesthetic Education Anchored in AI Literacy Under the Embodiment of Intelligent Technology. China Educational Technology, 2024, (5):120-127.

[10] Miao, F. C. The AI Competency Framework for Students: Laying the Foundation for

Citizens' Literacy in an Intelligent Society. China Educational Technology, 2024, 11:1-12.
[11] Chiu, T. K., Ahmad, Z., Ismailov, M., et al. What Are Artificial Intelligence Literacy and Competency? A Comprehensive Framework to Support Them. Computers and Education Open, 2024, 6, 100171.