Asian Journal of Distance Education

AI Literacy Level of Students: A Scale Development Study

Taha Aydogmus, Yavuz Samur

Abstract: This study aimed to develop a valid and reliable scale to measure artificial intelligence literacy among high school students in Türkiye. By establishing a multidimensional measurement framework and providing robust psychometric evidence, the study sought to contribute an assessment tool capable of supporting research, educational practice, and evidence-based policy development in AI literacy. Based on an extensive literature review, an initial pool of 21 items was generated and evaluated by experts to ensure content validity. Following a pilot study with 211 students, the refined instrument was administered to 645 high school students. Exploratory factor analysis revealed a five-factor structure comprising 20 items, accounting for 63.3% of the total variance. The identified factors were Basic Knowledge and Awareness, Ethical and Safe Use, Application Skills, Critical Perspective, and Evaluation. Confirmatory factor analysis indicated good model fit (CFI = 0.94; RMSEA = 0.05; GFI = 0.93; RMR = 0.06; χ^2/df = 2.729). Reliability analyses demonstrated high internal consistency across the overall scale and its subscales. Significant differences in AI literacy scores were observed across gender, grade level, and patterns of AI use. Despite being limited to secondary school students in Türkiye, the scale provides a robust measurement tool for assessing AI literacy and offers empirical evidence to inform curriculum design, teacher education, and national digital literacy initiatives.

Keywords: artificial intelligence literacy, generative AI, scale development, high school students, educational technology, psychometric validation, secondary education, AI ethics, critical thinking, digital literacy

Highlights

What is already known about this topic:

- Artificial intelligence literacy is increasingly considered a core competency for secondary education.
- Existing AI literacy assessments are mostly adult-oriented or lack strong psychometric validation.
- There is limited empirical evidence on multidimensional AI literacy among high school students.

What this paper contributes:

- Developed a multidimensional AI literacy scale specifically for high school students in Türkiye.
- Identified a five-factor structure covering knowledge, ethical use, application, critical perspective, and evaluation.
- Validated the scale using EFA and CFA with a large sample (n=645), explaining 63.3% of total variance and demonstrating high overall and subscale reliability.

Implications for theory, practice and/or policy:

- Provides a theoretically grounded framework for conceptualizing AI literacy in secondary education.
- Enables educators to assess students' AI literacy levels and design targeted instructional interventions.

Published by Asian Society for Open and Distance Education (ASODE), Japan
ISSN 1347-9008 http://www.asianjde.com/

- Supports teacher education programs and policymakers in designing evidence-based strategies for ethical, critical, and applied AI competencies.

Introduction

Artificial intelligence (AI) technologies are increasingly embedded in everyday life and educational environments, influencing how individuals access information, make decisions, and interact with digital systems. As AI-driven applications become more visible in schools, students are expected not only to use these technologies but also to understand their basic functioning, limitations, and ethical implications. In this context, artificial intelligence literacy has emerged as a key competency for secondary education, as reflected in major international frameworks and policy initiatives (UNESCO, 2022; OECD, 2025; Zhang & Magerko, 2025).

Artificial intelligence literacy broadly refers to the ability to understand, evaluate, and responsibly use AI systems. Although recent studies emphasize its importance, existing assessments of AI literacy are largely developed for adult or higher education populations and tend to focus on specific dimensions rather than a comprehensive skill set (Wang et al., 2025).

Secondary education constitutes a critical period for the development of AI literacy, as students increasingly encounter AI-based tools in academic, social, and personal contexts. Measuring AI literacy at this stage is essential for informing curriculum development, teacher education, and policy initiatives aimed at promoting ethical, critical, and effective AI use. However, existing international frameworks emphasize curriculum development rather than validated measurement instruments for assessing AI literacy among high school students (UNESCO, 2022).

In response to this gap, the present study aims to develop a valid and reliable scale to measure artificial intelligence literacy among high school students in Türkiye. By establishing a multidimensional measurement framework and providing robust psychometric evidence, this study seeks to contribute an assessment tool that can support research, educational practice, and evidence-based policy development in AI literacy.

Literature

Artificial intelligence literacy has gained increasing attention within educational research as artificial intelligence technologies become embedded in learning environments and everyday digital practices. Beyond technical proficiency, scholars emphasize that AI literacy involves understanding how AI systems function, evaluating their outputs critically, and engaging with these systems in ethical and socially responsible ways. As a result, the literature has gradually shifted from viewing AI literacy as a narrow technical competence toward conceptualizing it as a multidimensional construct grounded in broader digital literacy and educational technology frameworks. Together, this body of literature outlines the theoretical foundations of artificial intelligence literacy, its key dimensions, and its relevance in secondary education, while also revealing important gaps in existing measurement approaches.

Conceptualizing Artificial Intelligence Literacy

The concept of artificial intelligence literacy has emerged from earlier discussions on digital literacy, media literacy, and computational thinking. Initial digital literacy frameworks focused primarily on technical and cognitive skills required to access and use digital tools effectively (Eshet-Alkalai, 2004). Over time, these frameworks expanded to include critical thinking, ethical awareness, and the ability to evaluate digital content, reflecting the increasing complexity of digital technologies (European Commission, 2022).

Within this evolving landscape, artificial intelligence literacy has been defined as individuals' capacity to understand the basic principles of AI systems, interpret and evaluate algorithmic outputs, and make informed decisions in AI-mediated contexts (Zhang & Magerko, 2025). International organizations such as UNESCO emphasize that AI literacy should not be limited to knowing how to use AI tools but should also encompass awareness of data use, transparency, bias, and ethical implications, particularly in educational settings (UNESCO, 2022). More recently, international organizations have moved toward developing comprehensive, operationalized frameworks specifically targeting primary and secondary education. Notably, OECD (2025) defines AI literacy as the technical knowledge, durable skills, and future-ready attitudes required to thrive in a world influenced by AI, encompassing four domains: engaging with, creating with, managing, and designing AI. The growing recognition of AI literacy as an essential competency has also been reflected in scholarly discourse within educational technology and distance education. Bozkurt and Sharma (2024) highlight that the emergence of generative AI represents not merely a technological shift but a fundamental transformation in how humans interact with knowledge systems, underscoring the need for learners to develop not only technical proficiency but also critical awareness of AI's epistemic and ethical dimensions.

Dimensions of Artificial Intelligence Literacy

A growing body of research conceptualizes artificial intelligence literacy as a multidimensional construct. Rather than representing a single skill domain, AI literacy is widely conceptualized as a set of interrelated dimensions that include foundational knowledge, application skills, ethical awareness, and critical evaluation (Zhang & Magerko, 2025). Studies in educational technology highlight that students' interactions with AI-based tools require not only procedural knowledge but also higher-order thinking skills such as evaluating reliability, recognizing bias, and understanding the limitations of automated systems (Zawacki-Richter et al., 2019). Ethical considerations such as data privacy, accountability, and responsible use are consistently identified as integral components of AI literacy, particularly for young learners who may be more vulnerable to algorithmic influence (Baker et al., 2019). A systematic review of AI literacy in K-12 education by Casal-Otero et al. (2023), drawing on 179 studies, identified two broad approaches: learning experience-based and theoretical perspective-based. Their findins underscore that while substantial progress has been made in designing AI literacy frameworks, there remains a persistent gap between conceptual models and empirically validated measurement instruments, particularly for secondary school populations. More recently, Ng et al. (2024) developed and validated the AI Literacy Questionnaire (AILQ) for secondary school students in Hong Kong, grounded in an affective, behavioural, cognitive, and ethical (ABCE) framework. This instrument represents a significant step forward in addressing measurement gaps at the secondary level, though it was developed within a specific East Asian educational context, which limits its direct applicability across diverse national settings such as Türkiye. These findings support the need for AI literacy frameworks and measurement tools that reflect the construct's multidimensional nature. While the multidimensional conceptualization of AI literacy has gained broad acceptance, emerging critical perspectives in the literature invite more nuanced reflection. Pangrazio (2026) questions whether the literacy framework is itself an adequate response to the challenges posed by AI, arguing that current approaches risk becoming overly instrumental, focused on how to use AI rather than on critically examining the social, political, and economic systems within which AI operates. This critique does not undermine the value of developing AI literacy, but it underscores the importance of ensuring that literacy frameworks, including measurement tools derived from them, incorporate genuinely critical and reflective dimensions rather than defaulting to a purely skills-based orientation.

Artificial Intelligence Literacy in Secondary Education

Secondary education represents a critical stage for the development of artificial intelligence literacy. During this period, students increasingly encounter AI-driven applications in both formal learning contexts and everyday digital environments, including recommendation systems, generative AI tools, and adaptive learning platforms. Research suggests that early exposure to AI concepts can foster more informed, critical, and ethical engagement with technology in later educational and professional contexts (UNESCO, 2022). Despite this importance, much of the existing research on AI literacy focuses on higher education students, preservice teachers, or adult learners. Studies addressing K–12 or secondary education contexts remain limited, and those that do exist often emphasize instructional interventions rather than systematic assessment of students' AI literacy levels (Karaoğlan Yılmaz & Yılmaz, 2023). This disproportion highlights the need for empirically grounded approaches to understanding how AI literacy manifests among high school students.

Gaps in the Literature and Measurement Approaches

Within the Turkish context, existing studies on artificial intelligence literacy have predominantly focused on university students or preservice teachers, with relatively limited attention given to secondary school populations (Elçiçek, 2024; Karaoğlan Yılmaz & Yılmaz, 2023). While these studies contribute valuable insights into attitudes and perceptions toward AI, they offer limited evidence regarding the systematic assessment of AI literacy as a multidimensional construct among high school students. The absence of validated measurement tools tailored to this educational level restricts empirical research and hinders efforts to inform curriculum development and instructional practices. At the international level, a similar pattern can be observed. Although interest in artificial intelligence literacy has grown substantially, existing measurement approaches remain fragmented. Many existing instruments focus on attitudes toward AI or are tailored to specific populations, rather than conceptualizing AI literacy as a multidimensional construct encompassing knowledge, skills, ethical awareness, and critical evaluation (Long & Magerko, 2020). A systematic review of 16 AI literacy scales by Lintner (2024), assessing their psychometric quality using the COSMIN framework, found that while most instruments demonstrated good structural validity and internal consistency, critical measurement properties such as cross-cultural validity, test-retest reliability, and measurement error remained largely untested. Critically, not a single scale reviewed had been validated for cross-cultural applicability, and instruments designed specifically for secondary school populations remained scarce. These findings directly inform the present study's approach, which prioritizes rigorous psychometric validation including both EFA and CFA on independent samples and an explicit test-retest reliability procedure. Moreover, several studies report partial or insufficient psychometric validation, raising concerns about the reliability and validity of these measures when applied across different contexts or age groups. International frameworks, including those proposed by UNESCO, primarily emphasize curriculum design and competency descriptions related to AI education, while offering limited guidance on standardized and psychometrically validated instruments for assessing student-level AI literacy (UNESCO, 2022). As a result, there is a clear gap in the literature regarding reliable and context-appropriate measurement tools capable of capturing the multidimensional nature of artificial intelligence literacy among secondary school students. The urgency of developing a measurement tool tailored to secondary school students becomes particularly evident when examined alongside recent trends in adolescent AI use. Evidence suggests that AI adoption among teenagers has accelerated at a pace that outstrips the development of corresponding educational guidance. Pew Research Center (2026) reported that 54% of U.S. teens aged 13 to 17 now use AI chatbots to assist with schoolwork, a figure that doubled within just two years. Notably, this trend is most pronounced among older students: those in 11th and 12th grades show the highest rates of AI use for academic tasks, indicating that the secondary school years represent a peak period of AI engagement. Yet increased use has not been accompanied by proportional growth in ethical awareness or critical judgment. The same survey found that 59% of teens believe AI-assisted cheating has become a routine feature of student life, while only 18% consider using AI to write essays to be acceptable a pattern that reveals a substantial disconnect between how adolescents use these tools and how they understand the ethical boundaries of that use (Pew Research Center, 2026). These findings point to a pressing need for instruments capable of assessing AI literacy in ways that are sensitive to the specific behavioral patterns, developmental characteristics, and educational contexts of secondary school learners. Without such tools, efforts to identify competency gaps, design targeted interventions, or evaluate educational initiatives remain constrained by the absence of robust empirical evidence at this level.

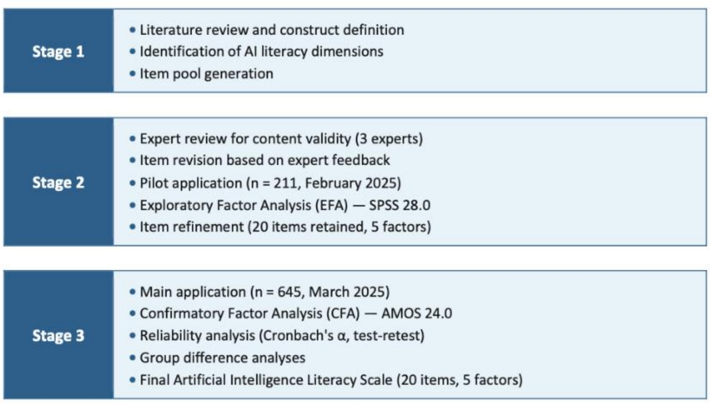
Methodology

In this study, a scale was developed to measure artificial intelligence literacy among high school students. The scale development process was grounded in the relevant literature on artificial intelligence literacy and educational measurement, drawing on theoretical frameworks and empirical studies to define the construct and its dimensions. Based on this conceptual foundation, an initial item pool was generated, expert opinions were obtained to ensure content validity, and a pilot study was conducted. Subsequently, the refined scale was administered to a larger sample, and the data were analyzed using exploratory and confirmatory factor analyses to establish the validity and reliability of the instrument.

Research Model/Design

This study employed a quantitative scale development design, a well-established methodological approach for constructing and validating psychological and educational measurement instruments (DeVellis, 2016; Hinkin, 1998). More specifically, the study adopted a cross-sectional survey design, in which data were collected at a single point in time from each study group to examine the psychometric properties of the developed instrument. Scale development research follows a systematic sequence of theoretically grounded and empirically rigorous phases, each of which builds upon the preceding one to ensure the resulting instrument is both conceptually sound and psychometrically robust. In the present study, this process was organized across five sequential phases: (1) construct definition through literature review, (2) item pool generation and expert review for content validity, (3) pilot application and exploratory factor analysis, (4) main application with confirmatory factor analysis and reliability analyses, and (5) finalization of the validated instrument. This procedure is illustrated in Figure 1. Each phase informed the subsequent one, ensuring that the theoretical grounding and psychometric integrity of the scale were established through multiple independent lines of evidence. A detailed overview of the pilot and main study applications is provided in Table 3.

Figure 1. The Stages of the Scale Development Process



Note. Scale development process informed by DeVellis (2016).

Data Collecting Tools

Data were collected using an online questionnaire consisting of two sections. The first section included demographic questions related to participants' gender, grade level, school type, and patterns of technology and AI use. The second section comprised the Artificial Intelligence Literacy Scale developed within the scope of this study. The initial item pool consisted of 21 items generated through a systematic review of the AI literacy literature, with particular reference to key theoretical frameworks that conceptualize AI literacy as a multidimensional construct encompassing knowledge, skills, ethical awareness, and critical evaluation (Long & Magerko, 2020; UNESCO, 2022; OECD, 2025; Zhang & Magerko, 2025). Items were developed to reflect these core dimensions and to be appropriate for the linguistic and cognitive level of secondary school students in Türkiye. The initial item pool was subsequently submitted for expert review. Three academic experts with specializations in educational measurement and evaluation, educational technology, statistics, and artificial intelligence reviewed the items for clarity, relevance, and content coverage. Based on their feedback, revisions were made to improve item quality and ensure alignment with the theoretical framework. After the pilot implementation, the scale was refined to a final structure consisting of 20 items grouped under five dimensions: Basic Knowledge and Awareness, Ethical and Safe Use, Application Skills, Critical Perspective, and Evaluation. All items were rated on a 5-point Likert scale ranging from 1 (strongly disagree) to 5 (strongly agree).

Sampling or Study Group

Participants were selected through convenience sampling, a non-probability technique in which individuals are included based on their accessibility and voluntary participation. Schools were identified based on practical accessibility, and student participation was entirely voluntary. Probability-based sampling methods such as simple random, stratified, cluster, or systematic sampling were not employed due to logistical constraints inherent in accessing secondary school populations. Nevertheless, to enhance sample diversity, students from different school types, grade levels, and geographic regions of Türkiye were deliberately included in the study. Two separate study groups were used during the scale development process.

Pilot Study Group: The pilot study group consisted of 211 high school students who participated voluntarily. This group was used to examine item performance and to explore the underlying factor structure of the scale through exploratory factor analysis.

Main Study Group: The main study group consisted of 645 high school students enrolled in public and private high schools in different regions of Türkiye. Participants represented various grade levels, including preparatory class and grades 9 through 12. Demographic characteristics of the participants are presented in Table 1.

Table 1. Demographic Characteristics of the Participants

Variable	Category	f	%
Gender	Male	273	42.3
	Female	372	57.7
School Sector	Public School	604	93.6
	Private School	41	6.4
School Type	Anatolian High School	542	84.0
	Science High School	32	5.0
	Vocational High School	26	4.0
	Imam Hatip High School	19	3.0
	Other (Vocational, Fine Arts, etc.)	26	4.0
Grade Level	Preparatory Class	1	<1.0
	9th Grade	227	35.0
	10th Grade	196	30.0
	11th Grade	136	21.0
	12th Grade	85	13.0

Note. Percentages are calculated based on the total sample size (N = 645).

Table 1 shows that, the sample consisted of 645 participants, of whom 57.7% were female and 42.3% were male. The majority of participants were enrolled in public schools (93.6%), while a smaller proportion attended private schools (6.4%). In terms of school type, most participants were studying at Anatolian high schools (84%), followed by science high schools (5%), vocational high schools (4%), and Imam Hatip high schools (3%). Regarding grade level, the largest groups were ninth-grade (35%) and tenth-grade students (30%), with fewer participants from upper grade levels.

Data Analysis

Data analyses were conducted using SPSS and AMOS software packages. Exploratory factor analysis (EFA) was performed on the pilot study data to identify the factor structure of the scale. Prior to EFA, sampling adequacy was assessed using the Kaiser–Meyer–Olkin (KMO) measure and Bartlett's Test of Sphericity. Principal Axis Factoring was used as the extraction method, as it is suited for identifying latent constructs from observed variables in scale development research (DeVellis, 2016). Given the theoretical expectation that the underlying factors would be interrelated, Direct Oblimin rotation was applied to allow correlations among factors. Factor retention was determined using the Kaiser criterion (eigenvalue > 1), supplemented by scree plot examination; both approaches supported the five-factor solution. Items with factor loadings below .30 were removed, and items exhibiting cross-loadings defined as substantial loadings on more than one factor were also excluded to ensure the clarity and interpretability of the factor structure. Confirmatory factor analysis (CFA) was conducted on the main study data to test the factor structure identified through EFA. As the dimensions of AI literacy are theoretically expected to be interrelated, a correlated factor model was specified and tested alongside an uncorrelated model. The correlated model demonstrated substantially better fit and was therefore retained as the final model, consistent with the multidimensional conceptualization of AI literacy in the literature (Long & Magerko, 2020; Zhang & Magerko, 2025). To avoid overfitting and to provide a more rigorous test of construct validity, EFA and CFA were conducted on independent samples. The factor structure was first explored using the pilot sample ($n = 211$) and subsequently confirmed using a separate main sample ($n = 645$), in line with best practices in scale development methodology (DeVellis, 2016). Model fit was evaluated using commonly reported fit indices, including χ^2/df, RMSEA, CFI, GFI, and RMR. Details regarding fit indices and acceptance criteria are presented in Table 2.

Table 2. Confirmatory Factor Analysis Fit Indices and Model Fit Results

Fit Index	Acceptable Threshold	Obtained Value
RMSEA	< .08	.052
CFI	> .90	.941
GFI	> .85	.937
TLI	> .90	.929
χ^2/df	< 5	2.729

Note. RMSEA = Root Mean Square Error of Approximation; CFI = Comparative Fit Index; GFI = Goodness of Fit Index; TLI = Tucker–Lewis Index.

Table 2 shows that, the results of the confirmatory factor analysis indicated that the proposed measurement model demonstrated an acceptable to good fit to the data. The χ^2/df ratio was below the recommended threshold of 5, and the RMSEA value indicated a good model fit. Additionally, the incremental fit indices (CFI and TLI) and the absolute fit index (GFI) exceeded the commonly accepted cutoff values, supporting the adequacy of the model.

Reliability and Validity

Content validity was ensured through a structured expert review process conducted prior to the pilot application. The draft item pool was submitted to three academic experts with specializations in educational technology, educational measurement and evaluation, statistics, and artificial intelligence. Experts were asked to evaluate each item in terms of its relevance to the construct, clarity of expression, and adequacy in representing the intended dimension of AI literacy. Based on their feedback, linguistic revisions were made to several items, wording was simplified where necessary, and certain items were reformulated to more accurately reflect the theoretical dimensions they were designed to measure. This process ensured that the final item pool was both conceptually aligned with the theoretical framework and accessible to secondary school students. Construct validity was subsequently examined through EFA and CFA procedures conducted on independent samples, ensuring that the factor structure identified in the pilot study was independently confirmed in the main study. Reliability was assessed following the establishment of construct validity, in accordance with standard scale development practice (DeVellis, 2016). Internal consistency was evaluated using Cronbach's alpha coefficients for the overall scale and each subdimension, and temporal stability was examined through test-retest procedures.

Research Procedures

Data collection was carried out through an online survey platform. Prior to participation, students were informed about the purpose of the study, voluntary participation, and confidentiality of responses. The scale development process followed these steps:

1. Review of the literature and definition of the AI literacy construct
2. Generation of the initial item pool
3. Expert review for content validity
4. Pilot application and item refinement
5. Main application and psychometric analyses

During the pilot study, the 21-item scale was administered to 211 high school students in February 2025. Item discrimination was examined through upper-lower group comparisons (27% criterion), and sampling adequacy was assessed prior to factor analysis. EFA identified a five-factor structure, and items failing to meet the factor loading threshold of .30 or exhibiting cross-loadings were removed, resulting in a refined 20-item scale for use in the main study.
An overview of the pilot and main applications is provided in Table 3.

Table 3. Overview of the Pilot and Main Study Applications

Analysis Type	Study Group	Number of Participants	Number of Items	Software Used
Exploratory Factor Analysis (EFA)	Pilot Study Group	211	21	SPSS 28.0
Confirmatory Factor Analysis (CFA)	Main Study Group	645	20	AMOS 24.0

Note. EFA = Exploratory Factor Analysis; CFA = Confirmatory Factor Analysis.

Table 3 shows that, the scale development process was conducted in two sequential phases: a pilot study and a main study. The pilot study was conducted in February 2025, and the main study was carried out in March 2025. Exploratory factor analysis was performed on the pilot study data obtained from 211 participants to explore the underlying factor structure of the scale. Subsequently, confirmatory factor analysis was conducted on the main study data collected from 645 participants to test the construct validity of the refined 20-item scale.

Findings

Consistent with scale development studies in educational technology and AI literacy research, the findings are presented sequentially, beginning with preliminary analyses, followed by exploratory and confirmatory factor analyses, reliability evidence, and group comparisons. In line with prior measurement studies, brief contextual references to existing literature are provided where appropriate, while detailed interpretation of the findings is reserved for the Discussion section.

Preliminary Analyses and Item Discrimination

In scale development research, preliminary item analysis serves a critical gatekeeping function: items that fail to discriminate between high- and low-scoring respondents, or datasets that violate the assumptions required for factor analysis, can compromise the structural validity of the resulting instrument. Examining item discrimination indices and assessing data suitability are therefore considered essential prerequisites for factor-analytic procedures (DeVellis, 2016; Tabachnick & Fidell, 2013). Accordingly, preliminary analyses were conducted to assess the discriminative power of the items and the appropriateness of the dataset for factor analysis. Item discrimination was examined using comparisons between the upper 27% and lower 27% groups based on total scale scores. The results indicated statistically significant differences between the upper and lower groups across all items, demonstrating that the items effectively differentiated students with higher and lower levels of artificial intelligence literacy.

Table 4. Upper 27%–Lower 27% Group Comparison Results for Item Discrimination

Group	n	Mean	SD	t	p	Mean Difference	95% CI
Upper Group	57	96.70	4.66	28.16	<.001	33.28	[30.93, 35.62]
Lower Group	57	63.42	7.60				

Note. SD = standard deviation; CI = confidence interval.

Table 4 shows that, the independent samples t-test revealed a statistically significant difference between the upper and lower groups. The mean score of the upper group (M = 96.70, SD = 4.66) was significantly higher than that of the lower group (M = 63.42, SD = 7.60), $t(112) = 28.16$, $p < .001$. The 95% confidence interval for the mean difference did not include zero, indicating strong discriminative power. Following item discrimination analysis, sampling adequacy and factorability were assessed. The Kaiser–Meyer–Olkin (KMO) measure of sampling adequacy exceeded the recommended threshold, and Bartlett's Test of Sphericity was statistically significant, indicating that the correlation matrix was suitable for factor analysis.

Table 5. Kaiser–Meyer–Olkin and Bartlett's Test of Sphericity Results

Measure	Value
Kaiser–Meyer–Olkin (KMO) Measure of Sampling Adequacy	.877
Bartlett's Test of Sphericity: χ^2	2075.54
Degrees of Freedom (df)	210
Significance (p)	<.001

Note. χ^2 = chi-square; df = degrees of freedom.

Table 5 shows that, the Kaiser–Meyer–Olkin (KMO) value was .877, indicating a high level of sampling adequacy for factor analysis. Bartlett's Test of Sphericity was statistically significant, $\chi^2\ (210) = 2075.54$, $p <.001$, demonstrating that the correlations among the items were sufficiently large for exploratory factor analysis. These results confirm that the dataset was suitable for factor extraction.

Exploratory Factor Analysis Findings

Exploratory factor analysis (EFA) is widely used in scale development and AI literacy research to uncover the latent factor structure of newly developed instruments, particularly when the dimensionality of a construct has not yet been empirically confirmed (Laupichler et al., 2023; Ng et al., 2024). In this study, EFA was conducted on the pilot sample (n=211) to explore the latent factor structure of the scale. The analysis resulted in a five-factor solution with eigenvalues greater than 1, explaining 63.293% of the total variance.

Table 6. Total Variance Explained by the Extracted Factors

Factor	Initial Eigenvalues: Total	Initial Eigenvalues: % of Variance	Initial Eigenvalues: Cumulative %	Extracted Eigenvalues: Total	Extracted Eigenvalues: % of Variance	Extracted Eigenvalues: Cumulative %
1	7.476	35.602	35.602	7.476	35.602	35.602
2	2.259	10.757	46.360	2.259	10.757	46.360
3	1.373	6.539	52.899	1.373	6.539	52.899

Factor	Initial Eigenvalues: Total	Initial Eigenvalues: % of Variance	Initial Eigenvalues: Cumulative %	Extracted Eigenvalues: Total	Extracted Eigenvalues: % of Variance	Extracted Eigenvalues: Cumulative %
4	1.151	5.483	58.382	1.151	5.483	58.382
5	1.031	4.912	63.293	1.031	4.912	63.293

Note. Only factors with eigenvalues greater than 1 are reported.

Table 6 shows that, the five-factor solution accounted for 63.293% of the total variance, with the first factor explaining 35.602% and the remaining four factors contributing additional explained variance ranging from 4.912% to 10.757%. Examination of factor loadings indicated that all retained items loaded satisfactorily on their respective factors without exhibiting problematic cross-loadings. Taken together, these findings suggest that the extracted factors represent a substantial proportion of the variance and support the multidimensional structure of the scale. Based on the conceptual content of the items, the five factors were labeled as Basic Knowledge and Awareness, Ethical and Safe Use, Application Skills, Critical Perspective, and Evaluation. To identify the factor loadings and subscale structure, the relevant analyses were conducted, and the results are presented in Table 7.

Table 7. Rotated Factor Loadings of the Artificial Intelligence Literacy Scale Items

Item	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5
Item 11	.778				
Item 1	.775				
Item 3	.707				
Item 8	.676				
Item 4	.668				
Item 13	.614			.305	
Item 9	.428				
Item 16		.846			
Item 17		.842			
Item 18		.729			
Item 19		.675			
Item 6			.886		
Item 7			.779		
Item 12			.570		
Item 10			.543		
Item 2			.497		.421
Item 21				.889	
Item 20				.878	
Item 5					.713
Item 14		.413			.582
Item 15					.388

Note. Factor loadings below .30 are suppressed. Values represent standardized loadings obtained after rotation.

As a result of the exploratory factor analysis, the initial item pool was reduced from 21 items to 20 items, forming the preliminary structure of the scale. Items with low factor loadings or substantial cross-loadings were removed to improve the clarity and interpretability of the factor structure. The remaining items demonstrated satisfactory loadings on their respective factors, supporting the proposed five-factor model.

Confirmatory Factor Analysis Findings

Confirmatory factor analysis is commonly employed in scale validation research to test whether a factor structure derived from exploratory analysis holds when applied to an independent dataset, thereby providing stronger evidence for construct validity (Byrne, 2016; Hair et al., 2019). In this study, CFA was conducted using the main sample (n = 645) to evaluate the five-factor model identified through EFA.

Initial CFA results indicated that the uncorrelated five-factor model did not provide adequate model fit. Therefore, a correlated five-factor model was specified and tested. The correlated model demonstrated improved fit indices and met commonly accepted criteria for model adequacy. The fit indices of the correlated model, including χ^2/df, RMSEA, CFI, GFI, and RMR, were within acceptable ranges, providing evidence in support of the proposed factor structure.

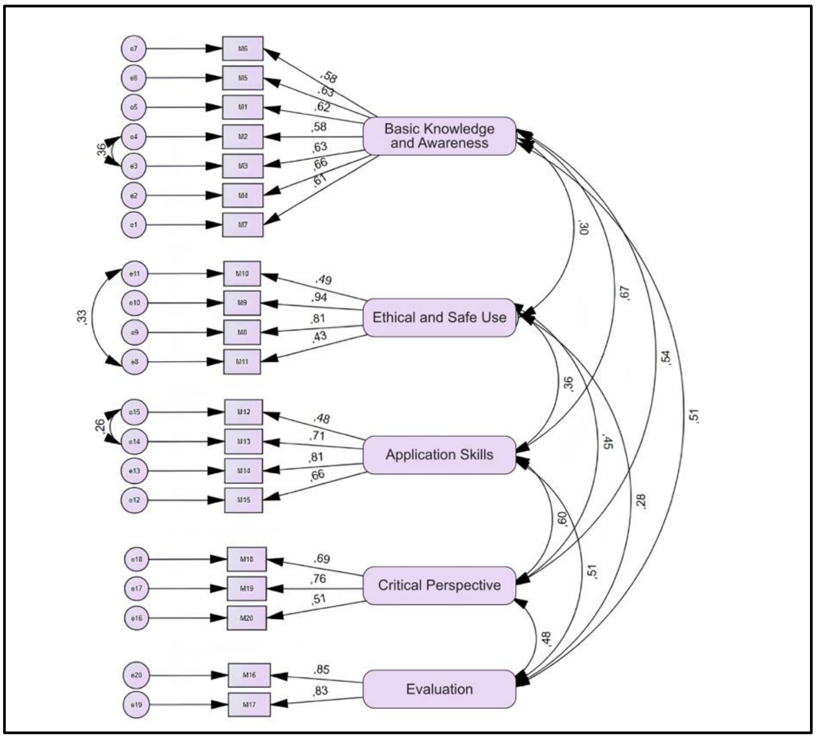
Table 8. Comparison of Model Fit Indices for Uncorrelated and Correlated CFA Models

Fit Index	Uncorrelated Model	Correlated Model
χ^2 (CMIN)	621.42	425.70
df	160	156
χ^2/df	3.88	2.73
RMR	.069	.060
GFI	.909	.937
RMSEA	.067	.052

Note. χ^2 = chi-square; df = degrees of freedom; RMR = Root Mean Square Residual; GFI = Goodness of Fit Index; RMSEA = Root Mean Square Error of Approximation.

Table 8 shows that, the correlated five-factor model demonstrated substantially better fit than the uncorrelated model. The χ^2/df ratio decreased from 3.88 to 2.73, indicating an improvement from acceptable to good model fit. Similarly, the RMSEA value decreased to .052, suggesting a close fit of the model to the data. Improvements were also observed in the absolute fit indices, with higher GFI and lower RMR values for the correlated model. Collectively, these results indicate that allowing correlations among the latent factors resulted in a more adequate representation of the data and provided strong support for the proposed five-factor structure.

Figure 2. Path Diagram of the Correlated Five-Factor CFA Model



As illustrated in Figure 2, all standardized factor loadings in the correlated five-factor CFA model were statistically significant and exceeded the commonly recommended threshold. The observed variables loaded strongly on their respective latent factors, providing evidence for the adequacy of the measurement model. In addition, the correlations among the latent factors were moderate and theoretically meaningful, supporting the multidimensional yet related nature of the construct. No further item removal or model modification was required, indicating that the final model demonstrated satisfactory construct representation and structural coherence.

Reliability Findings

Reliability evidence is a core component of scale validation studies. In this study, reliability analyses were conducted to examine both internal consistency and temporal stability of the Artificial Intelligence Literacy Scale. To provide a comprehensive account of the scale's psychometric properties, composite reliability (CR) and average variance extracted (AVE) were also calculated for each subscale alongside Cronbach's alpha coefficients. Internal consistency reliability was assessed using Cronbach's alpha, while CR values were computed using the formula proposed by Fornell and Larcker (1981) to provide an additional indicator of construct reliability. AVE values were calculated to assess the degree of convergent validity at the subscale level.

Table 9. Cronbach's Alpha Coefficients for the Total Scale and Subdimensions

Dimension	Cronbach's Alpha	Number of Items	CR	AVE
Factor 1: Basic Knowledge and Awareness	.871	7	.811	.380
Factor 2: Ethical and Safe Use	.806	4	.778	.491
Factor 3: Application Skills	.812	4	.765	.457
Factor 4: Critical Perspective	.784	2	.695	.438
Factor 5: Evaluation	.662	3	.827	.706
Total Scale	.906	20	-	-

Table 9 shows that, the scale demonstrated high internal consistency overall ($\alpha = .906$), with subscale Cronbach's alpha coefficients ranging from .662 to .871. CR values ranged from .695 to .827, all meeting or approaching the recommended .70 threshold, which supports the internal consistency of the subscales. AVE values ranged from .380 to .706. While AVE values for four of the five subscales fell below the commonly recommended .50 threshold, the CR values consistently exceeded this criterion, partially compensating for the lower AVE estimates. This pattern is not uncommon in multidimensional educational scales measuring complex and interrelated constructs and has been similarly observed in comparable AI literacy measurement studies (Laupichler et al., 2023; Ng et al., 2024). Temporal stability was examined using a test–retest procedure, in which the scale was administered twice to the same participants and Pearson correlation coefficients were calculated.

Table 10. Test–Retest Reliability Results for the Artificial Intelligence Literacy Scale

Item	Pearson's *r*
Item 1	.870
Item 2	.859
Item 3	.922
Item 4	.929
Item 5	.951
Item 6	.943
Item 7	.921
Item 8	.961
Item 9	.924
Item 10	.915
Item 11	.962
Item 12	.923
Item 13	.902
Item 14	.866
Item 15	.938

Item	Pearson's *r*
Item 16	.942
Item 17	.900
Item 18	.911
Item 19	.895
Item 20	.902
Total Score	.965

Note. All correlations are statistically significant (p < .001).

Table 10 shows that, test–retest reliability coefficients indicated strong temporal stability for both individual items and the overall scale. Item-level correlations ranged from high to very high, and the correlation for the total scale score was particularly strong (r = .967, p < .001). These findings demonstrate that the Artificial Intelligence Literacy Scale produces highly consistent scores across repeated administrations over time.

Group Difference Findings

Group comparisons are frequently incorporated into AI literacy scale validation studies to examine whether literacy competencies vary systematically across demographic characteristics such as gender and grade level, as well as technology usage patterns; findings that can inform targeted educational interventions (Ng et al., 2024; Wang et al., 2022).In this study, group difference analyses were conducted after establishing the validity and reliability of the scale.

Gender: An independent samples t-test was conducted to examine differences in AI literacy scores by gender. The results revealed statistically significant differences between male and female students on the overall scale and/or selected subdimensions.

Table 11. Independent Samples t-Test Results for AI Literacy Scores by Gender

Dimension	Gender	*n*	*M*	*SD*	Levene's Test (*F*, *p*)	*t* (*df*)	*p*
Factor 1: Basic Knowledge and Awareness	Male	273	24.06	6.06	8.54, .004	3.67 (533.57)	< .001
	Female	372	22.38	5.23			
Factor 2: Ethical and Safe Use	Male	273	15.69	3.72	10.58, .001	−4.16 (529.43)	< .001
	Female	372	16.82	3.17			
Factor 3: Application Skills	Male	273	15.75	3.59	0.46, .497	0.01 (643)	.995
	Female	372	15.75	3.66			
Factor 4: Critical Perspective	Male	273	7.27	2.46	0.19, .665	1.33 (643)	.185
	Female	372	7.01	2.48			
Factor 5: Evaluation	Male	273	12.40	2.37	1.22, .271	2.00 (643)	.046
	Female	372	12.02	2.33			

Note. M = mean; SD = standard deviation. Degrees of freedom were adjusted where Levene's test indicated unequal variances.

Table 11 shows that, statistically significant gender differences were observed for the overall AI literacy score and for selected subdimensions. Male students scored significantly higher than female students on Basic Knowledge and Awareness and on the total AI literacy score. In contrast, female students scored significantly higher on the Ethical and Safe Use dimension. No significant gender differences were found for Application Skills or Critical Perspective. A small but statistically significant difference was observed for the Evaluation dimension.

Level: Differences in AI literacy scores across grade levels were examined using one-way analysis of variance (ANOVA). The results indicated statistically significant differences among grade levels.

Table 12. One-Way ANOVA Results for AI Literacy Scores by Grade Level

Dimension	Source	SS	*df*	MS	*F*	*p*
Factor 1: Basic Knowledge and Awareness	Between Groups	112.67	4	28.17	0.88	.475
	Within Groups	20449.75	640	31.95		
	Total	20562.42	644			
Factor 2: Ethical and Safe Use	Between Groups	94.57	4	23.64	1.99	.095

Dimension	Source	SS	*df*	MS	*F*	*p*
	Within Groups	7605.76	640	11.88		
	Total	7700.32	644			
Factor 3: Application Skills	Between Groups	14.80	4	3.70	0.28	.891
	Within Groups	8475.01	640	13.24		
	Total	8489.81	644			
Factor 4: Critical Perspective	Between Groups	68.41	4	17.10	2.83	.024
	Within Groups	3869.67	640	6.05		
	Total	3938.08	644			
Factor 5: Evaluation	Between Groups	9.48	4	2.37	0.43	.790
	Within Groups	3567.66	640	5.57		
	Total	3577.14	644			

Note. SS = sum of squares; MS = mean square. p values less than .05 indicate statistical significance.

Table 12 shows that, one-way ANOVA results revealed a statistically significant difference across grade levels only for the Critical Perspective dimension, $F(4, 640) = 2.83$, $p = .024$. No statistically significant differences were observed across grade levels for Basic Knowledge and Awareness, Ethical and Safe Use, Application Skills, or Evaluation. These findings suggest that grade level is associated with differences in students' critical perspectives toward artificial intelligence, while other dimensions of AI literacy remain relatively stable across grade levels. Following the one-way ANOVA results, a Games–Howell post-hoc test was conducted to identify specific grade-level differences in the Critical Perspective subdimension. This procedure was selected due to unequal group sizes and potential violations of the homogeneity of variance assumption.

Table 13. Games-Howell Post-Hoc Test Results for AI Literacy Scores by Grade Level

(I) Grade Level	(J) Grade Level	Mean Difference (I–J)	SE	*p*
9th Grade	10th Grade	−0.224	0.244	.797
	11th Grade	−0.668	0.268	.063
	12th Grade	−0.865	0.304	.025*
10th Grade	9th Grade	0.224	0.244	.797
	11th Grade	−0.445	0.262	.326
	12th Grade	−0.641	0.294	.142
11th Grade	9th Grade	0.668	0.268	.063
	10th Grade	0.445	0.262	.326
	12th Grade	−0.197	0.318	.926
12th Grade	9th Grade	0.865	0.304	.025*
	10th Grade	0.641	0.294	.142
	11th Grade	0.197	0.318	.926

Note. SE = standard error. $p < .05$ indicates statistical significance.

Table 13 shows that, the Games–Howell post-hoc analysis revealed a statistically significant difference in the Critical Perspective subdimension between 9th- and 12th-grade students ($p = .025$). Specifically, 12th-grade students demonstrated higher levels of critical perspective compared to 9th-grade students. No other pairwise comparisons reached statistical significance.

School Type: One-way ANOVA was conducted to examine differences in AI literacy scores across school types. The analysis revealed statistically significant differences in overall scale scores and specific subdimensions.

Table 14. One-Way ANOVA Results for AI Literacy Scores by School Type

Dimension	Source	SS	*df*	MS	*F*	*p*
Factor 1: Basic Knowledge and Awareness	Between Groups	567.53	7	81.08	2.58	.012*
	Within Groups	19994.89	637	31.39		
	Total	20562.42	644			
Factor 2: Ethical and Safe Use	Between Groups	256.10	7	36.59	3.13	.003*

Dimension	Source	SS	*df*	MS	*F*	*p*
	Within Groups	7444.22	637	11.69		
	Total	7700.32	644			
Factor 3: Application Skills	Between Groups	187.32	7	26.76	2.05	.047*
	Within Groups	8302.48	637	13.03		
	Total	8489.81	644			
Factor 4: Critical Perspective	Between Groups	52.04	7	7.43	1.22	.290
	Within Groups	3886.04	637	6.10		
	Total	3938.08	644			
Factor 5: Evaluation	Between Groups	62.84	7	8.98	1.63	.125
	Within Groups	3514.30	637	5.52		
	Total	3577.14	644			

Note. SS = sum of squares; MS = mean square. p values less than .05 indicate statistical significance.

Table 14 shows that, one-way ANOVA results revealed statistically significant differences across school types for Basic Knowledge and Awareness, Ethical and Safe Use, and Application Skills. No statistically significant differences were observed for the Critical Perspective or Evaluation dimensions. These findings suggest that school type is associated with variations in certain dimensions of AI literacy, while other dimensions remain relatively consistent across different school contexts.

Table 15. Tukey Analysis Results for the Basic Knowledge and Awareness Subdimension by School Type

School Type (I)	School Type (J)	Mean Difference (I–J)	SE	*p*
Anatolian High School	Science High School	−3.22	0.98	.018*
	Vocational High School	1.69	1.12	.741
	Imam Hatip High School	−1.17	1.22	.962
	Fine Arts High School	−1.08	3.97	1.000
	Open High School	−0.92	3.97	1.000
	Other	−7.58	3.97	.474
Science High School	Anatolian High School	3.22	0.98	.018*
	Vocational High School	4.91	1.45	.013*
	Imam Hatip High School	2.05	1.52	.829
	Fine Arts High School	2.14	4.07	.998
	Open High School	4.14	4.07	.950
	Other	−4.36	4.07	.937
Vocational High School	Science High School	−4.91	1.45	.013*
	Anatolian High School	−1.69	1.12	.741
Other comparisons	—	—	—	ns

Note. SE = standard error. p < .05 indicates statistical significance. Only statistically significant pairwise comparisons are highlighted; non-significant comparisons (ns) are retained for completeness but condensed for clarity.

Table 15 shows that, Tukey HSD post-hoc analysis revealed statistically significant differences between students attending Science High Schools and those attending Anatolian High Schools and Vocational High Schools. Specifically, Science High School students scored significantly higher than both Anatolian High School and Vocational High School students. No other pairwise comparisons reached statistical significance. Following the one-way ANOVA results, a Games–Howell post-hoc test was conducted to examine pairwise differences among school types in the Ethical and Safe Use subdimension.

Table 16. Tukey Analysis Results for the Ethical and Safe Use Subdimension by School Type

School Type (I)	School Type (J)	Mean Difference (I–J)	SE	*p*
Anatolian High School	Fine Arts High School	−3.49	1.45	<.001*
Science High School	Fine Arts High School	−4.71	0.49	
Vocational High School	Fine Arts High School	−6.12	0.92	
Imam Hatip High School	Fine Arts High School	−4.05	0.59	
Fine Arts High School	Anatolian High School	3.49	1.45	
	Science High School	4.71	0.49	

School Type (I)	School Type (J)	Mean Difference (I–J)	SE	p
	Vocational High School	6.12	0.92	
	Imam Hatip High School	4.05	0.59	

Note. SE = standard error. p < .05 indicates statistical significance. Only statistically significant pairwise comparisons are reported for clarity.

Table 16 shows that, the Games–Howell post-hoc analysis revealed statistically significant differences in the Ethical and Safe Use subdimension primarily involving Fine Arts High Schools. Students attending Fine Arts High Schools scored significantly higher than students from Anatolian, Science, Vocational, and Imam Hatip High Schools. No other pairwise comparisons reached statistical significance.

Table 17. Games–Howell Post-Hoc Test Results for the Application Skills Subdimension Across School Types

School Type (I)	School Type (J)	Mean Difference (I–J)	SE	p
Anatolian High School	Science High School	−1.38	0.41	.022*
Science High School	Anatolian High School	1.38	0.41	.022*
	Vocational High School	3.15	0.89	.017*
Vocational High School	Science High School	−3.15	0.89	.017*

Note. SE = standard error. p<.05 indicates statistical significance. Only statistically significant pairwise comparisons are reported for clarity.

Table 17 shows that, the Games–Howell post-hoc analysis revealed statistically significant differences in the Application Skills subdimension involving Science High Schools. Specifically, students attending Science High Schools scored significantly higher than those attending Anatolian High Schools and Vocational High Schools. No other pairwise comparisons reached statistical significance.

Artificial Intelligence Use Patterns

Differences in AI literacy scores based on students' reported AI-related study and usage behaviors were examined using independent samples t tests and one-way ANOVA procedures, depending on the number of comparison groups.

Content Development Using AI: To examine whether AI literacy differed by students' self-reported use of AI for content development, descriptive statistics were calculated for each subdimension and the total score. Group differences were then tested using independent samples t tests.

Table 18. Descriptive Statistics by AI-based Content Development Status

Dimension	AI-Based Content Development	n	M	SD
Factor 1: Basic Knowledge and Awareness	Used AI for content development	144	25.91	5.33
	Did not use AI for content development	501	22.28	5.48
Factor 2: Ethical and Safe Use	Used AI for content development	144	16.24	3.30
	Did not use AI for content development	501	16.37	3.50
Factor 3: Application Skills	Used AI for content development	144	16.81	2.93
	Did not use AI for content development	501	15.45	3.76
Factor 4: Critical Perspective	Used AI for content development	144	7.67	2.23
	Did not use AI for content development	501	6.97	2.52
Factor 5: Evaluation	Used AI for content development	144	12.63	2.25
	Did not use AI for content development	501	12.05	2.37
Total AI Literacy Score	Used AI for content development	144	79.24	11.74
	Did not use AI for content development	501	73.12	12.40

Note. M = mean; SD = standard deviation. Values are reported for students who reported using AI to develop content versus those who did not.

Table 18 shows that, students who reported developing content using AI tended to have higher mean scores on Basic Knowledge and Awareness, Application Skills, Critical Perspective, Evaluation, and the total AI literacy score. Mean scores were similar across groups for Ethical and Safe Use. Assumptions were evaluated using Levene's test. Where the homogeneity of variance assumption was violated, Welch's t test results were reported.

Table 19. Levene's Test for Equality of Variances and Independent Samples t-Test Results by AI-Based Content Development Status

Dimension	Variance Assumption	Levene's F	Levene's p	t	df	p (2-tailed)	Mean Difference
Factor 1: Basic Knowledge and Awareness	Equal variances assumed	0.04	.834	7.04	643	<.001	3.63
	Equal variances not assumed	—	—	7.15	236.77	<.001	3.63
Factor 2: Ethical and Safe Use	Equal variances assumed	0.70	.403	−0.40	643	.689	−0.13
	Equal variances not assumed	—	—	−0.41	243.03	.679	−0.13
Factor 3: Application Skills	Equal variances assumed	20.95	p < .001	3.99	643	<.001	1.35
	Equal variances not assumed	—	—	4.58	291.91	<.001	1.35
Factor 4: Critical Perspective	Equal variances assumed	2.99	.084	3.01	643	.003	0.70
	Equal variances not assumed	—	—	3.22	257.86	.001	0.70
Factor 5: Evaluation	Equal variances assumed	0.47	.494	2.58	643	.010	0.57
	Equal variances not assumed	—	—	2.66	242.05	.008	0.57
Total AI Literacy Score	Equal variances assumed	1.44	.231	5.28	643	<.001	6.12
	Equal variances not assumed	—	—	5.45	242.53	<.001	6.12

Note. Degrees of freedom were adjusted where Levene's test indicated violation of the homogeneity of variance assumption.

Table 19 shows that, the homogeneity of variance assumption was met for Basic Knowledge and Awareness, Ethical and Safe Use, Critical Perspective, Evaluation, and the total score (Levene's $p > .05$), but was violated for Application Skills (Levene's $p < .001$). Independent samples t tests indicated that students who used AI for content development scored significantly higher on Basic Knowledge and Awareness, Application Skills (Welch-adjusted), Critical Perspective, Evaluation, and the total AI literacy score (all $p < .05$). No statistically significant difference was observed for Ethical and Safe Use ($p > .05$).

Daily Digital Device Use: Differences in AI literacy across daily digital device use levels were examined using one-way ANOVA. Prior to ANOVA, the homogeneity of variance assumption was assessed using Levene's test. Descriptive statistics for AI literacy subdimensions across daily digital device use durations are presented in Table 20.

Table 20. Means and Standard Deviations of AI Literacy Subdimensions by Daily Digital Device Use Duration

Dimension	Daily Digital Device Use	n	M	SD
Factor 1: Basic Knowledge and Awareness	About 1 hour	36	22.14	4.98
	1–3 hours	230	23.10	5.73
	3–5 hours	322	22.96	5.52
	More than 5 hours	57	24.40	6.36
	Total	645	23.09	5.65
Factor 2: Ethical and Safe Use	About 1 hour	36	17.36	3.51
	1–3 hours	230	16.92	2.94
	3–5 hours	322	16.05	3.54
	More than 5 hours	57	14.95	4.27
	Total	645	16.34	3.46
Factor 3: Application Skills	About 1 hour	36	14.86	3.97
	1–3 hours	230	15.70	3.51

Dimension	Daily Digital Device Use	n	M	SD
	3–5 hours	322	15.79	3.60
	More than 5 hours	57	16.35	4.06
	Total	645	15.75	3.63
Factor 4: Critical Perspective	About 1 hour	36	7.06	2.50
	1–3 hours	230	7.12	2.47
	3–5 hours	322	7.06	2.50
	More than 5 hours	57	7.33	2.53
	Total	645	7.12	2.47
Factor 5: Evaluation	About 1 hour	36	12.21	2.73
	1–3 hours	230	12.18	2.36
	3–5 hours	322	12.08	2.27
	More than 5 hours	57	12.31	2.29
	Total	645	12.18	2.36
Total AI Literacy Score	About 1 hour	36	73.81	11.91
	1–3 hours	230	75.17	11.86
	3–5 hours	322	73.94	12.48
	More than 5 hours	57	75.25	15.49
	Total	645	74.49	12.52

Note. M = mean; SD = standard deviation. Daily digital device use categories were grouped as about 1 hour, 1–3 hours, 3–5 hours, and more than 5 hours.

Table 20 shows that, descriptive trends suggest higher Basic Knowledge and Awareness and Application Skills scores among students reporting more than 5 hours of daily digital device use, whereas Ethical and Safe Use scores tended to decrease as usage duration increased. Other subdimensions exhibited relatively stable mean values across usage groups. Prior to conducting one-way ANOVA analyses, the assumption of homogeneity of variances was examined using Levene’s test for each AI literacy subdimension and the total scale score. The results of Levene’s test are presented in Table 21.

Table 21. Levene’s Test of Homogeneity of Variances for AI Literacy Scores by Daily Digital Device Use Duration

Dimension	F	p
Factor 1: Basic Knowledge and Awareness	2.019	.110
Factor 2: Ethical and Safe Use	4.056	.007
Factor 3: Application Skills	0.375	.771
Factor 4: Critical Perspective	0.425	.735
Factor 5: Evaluation	0.515	.672
Total AI Literacy Score	1.096	.350

Note. Levene’s test was used to assess the assumption of homogeneity of variances. p values greater than .05 indicate that the homogeneity assumption is met.

Table 21 shows that, the homogeneity assumption was met for all subdimensions except Ethical and Safe Use (p = .007), for which a Games–Howell post-hoc test was applied. One-way ANOVA results the scale are presented in Table 22.

Table 22. One-Way ANOVA Results for AI Literacy Scores by Daily Digital Device Use Duration

Dimension	Source	SS	df	MS	F	p
Factor 1: Basic Knowledge and Awareness	Between Groups	136.141	3	45.38	1.42	.235
	Within Groups	20,426.278	641	31.87		
	Total	20,562.419	644			
Factor 2: Ethical and Safe Use	Between Groups	252.478	3	84.16	7.24	<.001
	Within Groups	7,478.410	641	11.62		
	Total	7,700.319	644			
Factor 3: Application Skills	Between Groups	50.181	3	16.73	1.27	.284
	Within Groups	8,439.623	641	13.17		

Dimension	Source	SS	df	MS	F	p
	Total	8,489.805	644			
Factor 4: Critical Perspective	Between Groups	4.355	3	1.45		
	Within Groups	3,933.722	641	6.14	0.24	.871
	Total	3,938.078	644			
Factor 5: Evaluation	Between Groups	7.506	3	2.50		
	Within Groups	3,569.632	641	5.57	0.45	.718
	Total	3,577.138	644			

Note. SS = sum of squares; MS = mean square. p values less than .05 indicate statistical significance.

Table 22 shows that, one-way ANOVA results revealed a statistically significant difference across daily digital device use levels only for the Ethical and Safe Use subdimension, $F(3, 641) = 7.24$, $p < .001$. No statistically significant differences were observed for Basic Knowledge and Awareness, Application Skills, Critical Perspective, or Evaluation ($p > .05$). These findings indicate that the amount of daily digital device use is associated with differences in students' ethical and safe use of artificial intelligence, whereas other dimensions of AI literacy appear to remain relatively stable across different levels of digital device use. Because the homogeneity of variance assumption was violated for the Ethical and Safe Use subdimension, a Games–Howell post hoc test was conducted to examine pairwise differences across daily digital device use durations.

Table 23. Games–Howell Post Hoc Test Results for the Ethical and Safe Use Subdimension by Daily Digital Device Use Duration

I (Daily Use Duration)	J (Daily Use Duration)	Mean Difference (I–J)	SE	p
About 1 hour	1–3 hours	0.44	0.62	.891
	3–5 hours	1.31	0.62	.163
	More than 5 hours	2.41	0.81	.020*
1–3 hours	About 1 hour	−0.44	0.62	.891
	3–5 hours	0.87	0.28	.010*
	More than 5 hours	1.97	0.60	.008*
3–5 hours	About 1 hour	−1.31	0.62	.163
	1–3 hours	−0.87	0.28	.010*
	More than 5 hours	1.11	0.60	.261
More than 5 hours	About 1 hour	−2.41	0.81	.020*
	1–3 hours	−1.97	0.60	.008*
	3–5 hours	−1.11	0.60	.261

Note. SE = standard error. $p < .05$ indicates statistical significance.

Table 23 shows that, the Games–Howell post hoc analysis indicated statistically significant differences in the Ethical and Safe Use subdimension across daily digital device use levels. Students reporting more than five hours of daily device use scored significantly lower than those using devices for about one hour ($p = .020$) and 1–3 hours per day ($p = .008$). Additionally, students in the 3–5 hours group scored significantly lower than those in the 1–3 hours group ($p = .010$). Overall, these findings suggest that extended daily digital device use is associated with lower levels of ethical and safe artificial intelligence practices. While frequent digital engagement may increase technical familiarity, it does not necessarily support responsible or reflective AI use, underscoring the need for explicit ethical guidance within AI literacy education.

Daily AI Use Hours: Differences in AI literacy across daily AI use time levels were examined using one-way ANOVA. Descriptive statistics were computed for each subdimension and the total score. Levene's test was used to evaluate the homogeneity of variances prior to inferential analyses. Descriptive statistics were calculated to examine differences in AI literacy scores across levels of daily artificial intelligence use. Mean scores and standard deviations for each subdimension and the total scale are presented in Table 24.

Table 24. Means and Standard Deviations of AI Literacy Subdimensions by Daily Artificial Intelligence Use Duration

Dimension	Daily AI Use Duration	n	M	SD
Factor 1: Basic Knowledge and Awareness	About 1 hour	131	20.86	5.52

Dimension	Daily AI Use Duration	n	M	SD
	1–3 hours	425	23.27	5.40
	3–5 hours	69	25.39	6.37
	More than 5 hours	16	25.25	6.60
Factor 2: Ethical and Safe Use	About 1 hour	131	16.49	3.63
	1–3 hours	425	16.32	3.35
	3–5 hours	69	16.22	3.54
	More than 5 hours	16	15.50	4.69
Factor 3: Application Skills	About 1 hour	131	14.51	4.13
	1–3 hours	425	15.98	3.54
	3–5 hours	69	16.55	3.48
	More than 5 hours	16	15.81	4.90
Factor 4: Critical Perspective	About 1 hour	131	6.75	2.69
	1–3 hours	425	7.27	2.35
	3–5 hours	69	7.06	2.52
	More than 5 hours	16	6.63	3.12
Factor 5: Evaluation	About 1 hour	131	12.15	2.40
	1–3 hours	425	12.14	2.30
	3–5 hours	69	12.58	2.15
	More than 5 hours	16	11.25	3.87
Total AI Literacy Score	About 1 hour	131	70.76	12.63
	1–3 hours	425	74.99	11.81
	3–5 hours	69	77.80	13.23
	More than 5 hours	16	74.44	18.94

Table 24 shows that, mean scores for Basic Knowledge and Awareness increased with longer daily AI use durations, with the highest scores observed among students reporting 3–5 hours and more than 5 hours of use per day. Application Skills scores followed a similar pattern, whereas Ethical and Safe Use scores showed a slight decreasing trend at higher usage levels. Critical Perspective and Evaluation subdimensions exhibited relatively stable mean values across usage groups. Overall AI literacy scores were highest among students reporting 3–5 hours of daily AI use. Prior to conducting one-way analysis of variance, the assumption of homogeneity of variances was examined using Levene's test for each AI literacy subdimension and the total scale.

Table 25. Levene's Test of Homogeneity of Variance for Daily Artificial Intelligence Use Duration

Dimension	F	p
Factor 1: Basic Knowledge and Awareness	1.238	.293
Factor 2: Ethical and Safe Use	1.115	.308
Factor 3: Application Skills	1.115	.348
Factor 4: Critical Perspective	1.426	.224
Factor 5: Evaluation	1.204	.172
Total AI Literacy Score	1.604	.206

Note. p values greater than .05 indicate that the assumption of homogeneity of variances is met.

Table 25 shows that, Table 25 shows that the homogeneity assumption was satisfied for all subdimensions ($p > .05$), supporting the use of standard one-way ANOVA procedures.

Table 26. One-Way ANOVA Results for AI Literacy Scores by Daily Artificial Intelligence Use Duration

Dimension	Source	SS	df	MS	F	p
Factor 1: Basic Knowledge and Awareness	Between Groups	1248.40	4	312.10	10.34	< .001
	Within Groups	19314.02	640	30.18		
	Total	20562.42	644			
Factor 2: Ethical and Safe Use	Between Groups	43.73	4	10.93	0.91	.455
	Within Groups	7656.59	640	11.96		
	Total	7700.32	644			

Dimension	Source	SS	df	MS	F	p
Factor 3: Application Skills	Between Groups	288.68	4	72.17	5.63	< .001
	Within Groups	8201.13	640	12.81		
	Total	8489.81	644			
Factor 4: Critical Perspective	Between Groups	32.78	4	8.20	1.34	.252
	Within Groups	3905.29	640	6.10		
	Total	3938.08	644			
Factor 5: Evaluation	Between Groups	38.85	4	9.71	1.76	.136
	Within Groups	3538.29	640	5.53		
	Total	3577.14	644			

Note. SS = sum of squares; MS = mean square. p values less than .05 indicate statistical significance.

Table 26 shows that, one-way ANOVA results indicated statistically significant differences across daily AI use levels for Basic Knowledge and Awareness, $F(4, 640) = 10.34$, $p < .001$, and Application Skills, $F(4, 640) = 5.63$, $p < .001$. In contrast, no statistically significant differences were observed for Ethical and Safe Use, Critical Perspective, or Evaluation ($p > .05$). These findings suggest that increased daily engagement with artificial intelligence is associated with higher levels of basic AI knowledge and application-related skills, while ethical awareness, critical perspective, and evaluative competencies remain relatively stable across usage levels. Following the significant one-way ANOVA result for the Basic Knowledge and Awareness dimension, Tukey's Honestly Significant Difference (HSD) post-hoc test was conducted to identify specific differences between daily AI use groups.

Table 27. Tukey HSD Post-Hoc Test Results for the Basic Knowledge and Awareness Subdimension by Daily AI Use Level

(I) Daily AI Use	(J) Daily AI Use	Mean Difference (I–J)	SE	p
~1 hour	1–3 hours	−2.42	0.55	<.001
	3–5 hours	−4.54	0.82	<.001
	>5 hours	−4.40	1.45	.022
1–3 hours	~1 hour	2.42	0.55	<.001
	3–5 hours	2.42	0.55	<.001
	>5 hours	−2.12	0.71	.026
3–5 hours	~1 hour	1.98	1.40	.619
	1–3 hours	−5.73	2.76	.232
	>5 hours	4.54	0.82	<.001
>5 hours	~1 hour	2.12	0.71	.026
	1–3 hours	0.14	1.52	1.000
	3–5 hours	−3.61	2.83	.705

Note. SE = standard error. p values less than .05 indicate statistical significance.

Table 27 shows that, Tukey HSD post-hoc analysis revealed multiple statistically significant pairwise differences in the Basic Knowledge and Awareness subdimension. Students who reported using artificial intelligence for approximately one hour per day scored significantly lower than those using AI for 1–3 hours, 3–5 hours, and more than 5 hours per day. In addition, students in the 1–3 hour usage group scored significantly higher than those in the approximately one-hour group and lower than those in the 3–5 hour usage group. These findings indicate a clear positive association between increased daily AI use and higher levels of basic AI knowledge and awareness, with particularly pronounced differences between low and moderate-to-high usage groups. Following the significant one-way ANOVA result for the Application Skills subdimension, Tukey's Honestly Significant Difference post-hoc test was conducted to examine pairwise differences across daily artificial intelligence use levels.

Table 28. Tukey HSD Post-Hoc Test Results for the Application Skills Subdimension by Daily AI Use Level

(I) Daily AI Use	(J) Daily AI Use	Mean Difference (I–J)	SE	p
~1 hour	1–3 hours	−1.47	0.36	< .001
	3–5 hours	−2.04	0.53	.001
	>5 hours	−1.30	0.95	.646

(I) Daily AI Use	(J) Daily AI Use	Mean Difference (I–J)	SE	p
	~1 hour	1.47	0.36	< .001
1–3 hours	3–5 hours	1.47	0.36	< .001
	>5 hours	−0.57	0.46	.739
	~1 hour	0.17	0.91	1.000
3–5 hours	1–3 hours	−2.02	1.80	.795
	>5 hours	2.04	0.53	.001
	~1 hour	0.57	0.46	.739
>5 hours	1–3 hours	0.74	0.99	.946
	3–5 hours	−1.45	1.84	.934

Note. SE = standard error. p values less than .05 indicate statistical significance.

Table 28 shows that, the Tukey HSD post-hoc analysis revealed several statistically significant differences in the Application Skills subdimension. Students who reported using artificial intelligence for approximately one hour per day scored significantly lower than those using AI for 1–3 hours and 3–5 hours per day. In addition, students in the 1–3 hour usage group demonstrated significantly higher application skills compared to those in the approximately one-hour group. A significant difference was also observed between the 3–5 hour and more than 5 hour usage groups, favoring the 3–5 hour group. No other pairwise comparisons reached statistical significance. Overall, these findings suggest that moderate levels of daily AI use are associated with higher application skills, whereas very low or very high usage does not consistently correspond to higher performance.

Number of AI Applications Used: To examine whether AI literacy differed by the number of AI applications used, descriptive statistics were computed across usage-count categories, followed by one-way ANOVA analyses. Participants were grouped based on the total number of different AI tools they reported using, ranging from one application to more than five applications. Mean scores across all AI literacy subdimensions generally increased as the number of AI applications used increased. Students who reported using a greater variety of AI applications demonstrated higher levels of basic knowledge and awareness, application skills, critical perspective, and evaluation. This upward trend was particularly pronounced in the Application Skills and Evaluation subdimensions, suggesting that exposure to multiple AI tools may contribute to more advanced and diversified AI literacy competencies. However, descriptive patterns alone do not indicate statistical significance, and these findings were further examined using inferential analyses in subsequent sections. Prior to conducting one-way analysis of variance (ANOVA), the assumption of homogeneity of variances was examined using Levene's test for each AI literacy subdimension and the total scale score across groups defined by the number of AI applications used.

Table 29. Levene's Test of Homogeneity of Variances by Number of AI Applications Used

Dimension	F	p
Factor 1: Basic Knowledge and Awareness	1.158	.328
Factor 2: Ethical and Safe Use	0.183	.969
Factor 3: Application Skills	6.159	<.001
Factor 4: Critical Perspective	1.654	.144
Factor 5: Evaluation	2.952	.012
Total AI Literacy Score	0.657	.656

Note. A non-significant Levene's test ($p > .05$) indicates that the assumption of homogeneity of variances is met.

Table 29 shows that, Levene's test results indicated that the assumption of homogeneity of variances was satisfied for Basic Knowledge and Awareness, Ethical and Safe Use, Critical Perspective, and the total AI literacy score ($p > .05$). However, the assumption was violated for the Application Skills and Evaluation subdimensions ($p < .05$). Accordingly, standard one-way ANOVA procedures were applied for dimensions meeting the homogeneity assumption, whereas more robust post-hoc procedures appropriate for unequal variances were employed for dimensions in which this assumption was not met. Following the examination of variance homogeneity, one-way analysis of variance (ANOVA) was conducted to investigate differences in AI literacy subdimensions based on the number of artificial intelligence applications used by students.

Table 30. One-Way ANOVA Results by Number of AI Applications Used

Dimension	Source	SS	df	MS	F	p
Factor 1: Basic Knowledge and Awareness	Between Groups	3406.08	5	681.22	25.37	<.001
	Within Groups	17156.34	639	26.85		
	Total	20562.42	644			
Factor 2: Ethical and Safe Use	Between Groups	70.08	5	14.02		
	Within Groups	7630.24	639	11.94	1.17	.320
	Total	7700.32	644			
Factor 3: Application Skills	Between Groups	630.50	5	126.30		
	Within Groups	7958.31	639	12.45	8.54	<.001
	Total	8489.81	644			
Factor 4: Critical Perspective	Between Groups	134.84	5	26.97		
	Within Groups	3803.57	639	5.95	4.52	<.001
	Total	3938.08	644			
Factor 5: Evaluation	Between Groups	134.84	5	26.97		
	Within Groups	3442.30	639	5.39	5.01	<.001
	Total	3577.14	644			

Note. SS = sum of squares; MS = mean square. p values less than .05 indicate statistical significance.

Table 30 shows that, the one-way ANOVA results revealed statistically significant differences across groups for Basic Knowledge and Awareness, Application Skills, Critical Perspective, and Evaluation ($p < .001$). In contrast, no significant differences were observed for the Ethical and Safe Use subdimension ($p = .320$). These findings indicate that students' AI literacy levels in several subdimensions vary according to the number of AI applications they use. Consistent with the Levene's test results, appropriate post-hoc analyses were conducted to further examine the specific group differences for dimensions exhibiting significant main effects. Following the significant one-way ANOVA result for the Basic Knowledge and Awareness subdimension, Tukey's Honestly Significant Difference post-hoc test was conducted to identify specific group differences based on the number of artificial intelligence applications used.

Table 31. Tukey HSD Post-Hoc Test Results for the Basic Knowledge and Awareness Subdimension by Number of AI Applications Used

(I) Number of AI Applications	(J) Number of AI Applications	Mean Difference (I–J)	SE	p
1 application	2 applications	−2.77	0.48	<.001
	3 applications	−4.01	0.73	<.001
	4 applications	−6.11	1.19	<.001
	5 applications	−10.43	1.52	<.001
	More than 5 applications	−9.29	1.75	<.001
2 applications	1 application	2.77	0.48	<.001
	3 applications	−1.24	0.79	.615
	4 applications	−3.34	1.22	.071
	5 applications	−7.66	1.55	<.001
	More than 5 applications	−6.52	1.77	.003
3 applications	1 application	4.01	0.73	<.001
	2 applications	1.24	0.79	.615
	4 applications	−2.10	1.34	.623
	5 applications	−6.42	1.64	.001
	More than 5 applications	−5.28	1.86	.052
4 applications	1 application	6.11	1.19	<.001
	2 applications	3.34	1.22	.071
	3 applications	2.10	1.34	.623
	5 applications	−4.32	1.89	.203

(I) Number of AI Applications	(J) Number of AI Applications	Mean Difference (I–J)	SE	p
	More than 5 applications	−3.18	2.08	.646
5 applications	1 application	10.43	1.52	<.001
	2 applications	7.66	1.55	<.001
	3 applications	6.42	1.64	.001
	4 applications	4.32	1.89	.203
	proficient than 5 applications	1.14	2.28	.996
More than 5 applications	1 application	9.29	1.75	<.001
	2 applications	6.52	1.77	.003
	3 applications	5.28	1.86	.052
	4 applications	3.18	2.08	.646
	5 applications	−1.14	2.28	.996

Note. SE = standard error. p values less than .05 indicate statistically significant differences.

Table 31 shows that, the Tukey HSD post-hoc analysis revealed a clear and systematic pattern of differences across groups. Students who reported using a greater number of AI applications demonstrated significantly higher basic knowledge and awareness compared to those using fewer applications. In particular, students using five or more AI applications scored significantly higher than those using one, two, or three applications. Differences between adjacent usage groups (e.g., two vs. three applications) were generally not statistically significant. Overall, these findings suggest a positive association between the breadth of AI application use and foundational AI literacy. Given the significant main effect of the number of AI applications used on the Application Skills subdimension, Tukey's Honestly Significant Difference (HSD) post-hoc test was conducted to examine pairwise group differences.

Table 32. Tukey HSD Post-Hoc Test Results for the Application Skills Subdimension by Number of AI Applications Used

(I) Number of AI Applications	(J) Number of AI Applications	Mean Difference (I–J)	SE	p
1 application	2 applications	−1.44	0.32	<.001
	3 applications	−1.37	0.47	.047
	4 applications	−2.64	0.53	<.001
	5 applications	−4.02	0.49	<.001
	More than 5 applications	−2.27	0.65	.052
2 applications	1 application	1.44	0.32	<.001
	3 applications	0.07	0.49	1.000
	4 applications	−1.20	0.56	.292
	5 applications	−2.58	0.52	.001
	More than 5 applications	−0.83	0.67	.811
3 applications	1 application	1.37	0.47	.047
	2 applications	−0.07	0.49	1.000
	4 applications	−1.27	0.65	.389
	5 applications	−2.65	0.62	.002
	More than 5 applications	−0.90	0.75	.832
4 applications	1 application	2.64	0.53	<.001
	2 applications	1.20	0.56	.292
	3 applications	1.27	0.65	.389
	5 applications	−1.38	0.67	.336
	More than 5 applications	0.37	0.80	.997
5 applications	1 application	4.02	0.49	<.001
	2 applications	2.58	0.52	.001
	3 applications	2.65	0.62	.002
	4 applications	1.38	0.67	.336
	More than 5 applications	1.75	0.77	.262

(I) Number of AI Applications	(J) Number of AI Applications	Mean Difference (I–J)	SE	p
More than 5 applications	1 application	2.27	0.65	.052
	2 applications	0.83	0.67	.811
	3 applications	0.90	0.75	.832
	4 applications	−0.37	0.80	.997
	5 applications	−1.75	0.77	.810

Note. SE = standard error. p values less than .05 indicate statistically significant differences.

Table 32 shows that, the Tukey HSD post-hoc analysis indicated that students who reported using a greater number of AI applications generally demonstrated higher application skills. Specifically, students using four or five AI applications scored significantly higher than those using one application. Similarly, the five-application group showed significantly higher scores compared to both the two- and three-application groups. In contrast, differences involving the group using more than five applications were not consistently significant, likely due to the small sample size in this category. Overall, the findings suggest that increased exposure to multiple AI tools is associated with enhanced application skills, with the strongest differences observed between low and moderate-to-high usage groups. Following the significant one-way ANOVA result for the Critical Perspective subdimension, Tukey's Honestly Significant Difference (HSD) post-hoc test was conducted to identify specific group differences based on the number of artificial intelligence applications used.

Table 33. Tukey HSD Post-Hoc Test Results for the Critical Perspective Subdimension by Number of AI Applications Used

(I) Number of AI Applications	(J) Number of AI Applications	Mean Difference (I–J)	SE	p
1 application	2 applications	−0.50	0.23	.234
	3 applications	−0.69	0.34	.325
	4 applications	−1.02	0.49	.338
	5 applications	−2.01	0.59	.046
	More than 5 applications	−2.40	0.45	.004
2 applications	1 application	0.50	0.23	.234
	3 applications	−0.19	0.36	.995
	4 applications	−0.52	0.51	.908
	5 applications	−1.50	0.60	.195
	More than 5 applications	−1.89	0.47	.019
3 applications	1 application	0.69	0.34	.325
	2 applications	0.19	0.36	.995
	4 applications	−0.33	0.57	.991
	5 applications	−1.32	0.65	.373
	More than 5 applications	−1.70	0.53	.049
4 applications	1 application	1.02	0.49	.338
	2 applications	0.52	0.51	.908
	3 applications	0.33	0.57	.991
	5 applications	−0.98	0.75	.774
	More than 5 applications	−1.37	0.65	.307
5 applications	1 application	2.01	0.59	.046
	2 applications	1.50	0.60	.195
	3 applications	1.32	0.65	.373
	4 applications	0.98	0.75	.774
	More than 5 applications	−0.39	0.72	.994
More than 5 applications	1 application	2.40	0.45	.004
	2 applications	1.89	0.47	.019
	3 applications	1.70	0.53	.049
	4 applications	1.37	0.65	.307
	5 applications	0.39	0.72	.994

Note. SE = standard error. p values less than .05 indicate statistically significant differences.

Table 33 shows that, the Tukey HSD post-hoc analysis revealed that students who reported using five or more AI applications demonstrated significantly higher critical perspective scores compared to those using only one application. Similarly, students using more than five AI applications scored significantly higher than those using one, two, or three applications. In contrast, differences among groups using one to four applications were generally not statistically significant. These findings suggest that a broader engagement with multiple AI tools is associated with more advanced critical perspectives toward artificial intelligence, particularly when the number of applications used exceeds four. Following the significant one-way ANOVA result for the Evaluation subdimension, Tukey's Honestly Significant Difference (HSD) post-hoc test was conducted to examine pairwise differences across groups defined by the number of artificial intelligence applications used.

Table 34. Tukey HSD Post-Hoc Test Results for the Evaluation Subdimension by Number of AI Applications Used

(I) Number of AI Applications	(J) Number of AI Applications	Mean Difference (I–J)	SE	p
1 application	2 applications	−0.38	0.22	.494
	3 applications	−0.43	0.29	.428
	4 applications	−1.07	0.46	.357
	5 applications	−2.40	0.48	.008
	More than 5 applications	−2.29	0.30	<.001
2 applications	1 application	0.38	0.22	.494
	3 applications	−0.05	0.32	1.000
	4 applications	−0.69	0.48	.012
	5 applications	−2.02	0.50	<.001
	More than 5 applications	−1.91	0.33	.019
3 applications	1 application	0.43	0.29	.428
	2 applications	0.05	0.32	1.000
	4 applications	−0.64	0.52	.816
	5 applications	−1.97	0.37	.017
	More than 5 applications	−1.86	0.39	<.001
4 applications	1 application	1.07	0.46	.357
	2 applications	0.69	0.48	.699
	3 applications	0.64	0.52	.816
	5 applications	−1.33	0.64	.328
	More than 5 applications	−1.22	0.52	.212
5 applications	1 application	2.40	0.48	.003
	2 applications	2.02	0.50	.012
	3 applications	1.97	0.37	.017
	4 applications	1.33	0.64	.328
	More than 5 applications	0.11	0.52	.996
More than 5 applications	1 application	2.29	0.30	<.001
	2 applications	1.91	0.33	.019
	3 applications	1.86	0.39	<.001
	4 applications	1.22	0.52	.212
	5 applications	−0.11	0.52	1.000

Note. SE = standard error. p values less than .05 indicate statistically significant differences.

Table 34 shows that, the Tukey HSD post-hoc analysis indicated that students who reported using five or more AI applications scored significantly higher on the Evaluation subdimension compared to those using one, two, or three applications. Similarly, students using more than five AI applications demonstrated significantly higher evaluation scores than those using up to three applications. In contrast, differences among groups using one to four applications were generally not statistically significant. These findings suggest that a broader engagement with multiple AI tools is associated with more advanced evaluative competencies related to artificial intelligence.

In summary, the results provide comprehensive evidence for the psychometric robustness of the Artificial Intelligence Literacy Scale. The findings support a stable five-factor structure, confirmed through both exploratory and confirmatory factor analyses, alongside strong internal consistency and temporal stability. In addition, the scale demonstrated sensitivity in detecting meaningful differences across student groups based on demographic characteristics and patterns of artificial intelligence use, including frequency, intensity, and diversity of application engagement. These results establish the scale as a valid and reliable instrument for assessing artificial intelligence literacy among secondary school students. The implications of these findings are further examined in the Discussion section.

Discussion

Factor Structure

The findings of this study support a five-factor structure of artificial intelligence literacy among high school students, encompassing Basic Knowledge and Awareness, Ethical and Safe Use, Application Skills, Critical Perspective, and Evaluation. This multidimensional structure is consistent with contemporary conceptualizations of AI literacy, which emphasize that effective engagement with AI systems requires not only technical understanding but also ethical awareness, critical judgment, and evaluative reasoning (Zhang & Magerko, 2025). The exploratory factor analysis revealed that these five dimensions jointly explained a substantial proportion of the total variance, indicating that AI literacy cannot be adequately represented as a unidimensional construct in secondary education contexts. Similar multidimensional structures have been reported in prior AI literacy and digital competence frameworks, which highlight the interconnected yet distinct nature of knowledge, application, and ethical–critical dimensions (Zawacki-Richter et al., 2019; UNESCO, 2022). The present findings extend this literature by empirically validating such a structure specifically for high school students. Confirmatory factor analysis further supported the proposed factor structure, with the correlated five-factor model demonstrating acceptable model fit. The presence of correlations among factors is theoretically meaningful, as students' basic understanding of AI systems is closely related to their ability to apply these systems responsibly and to evaluate their outputs critically. This result aligns with previous scale development studies suggesting that AI-related competencies tend to co-develop rather than operate independently.

Validity and Reliability Evidence

The validity and reliability findings provide strong evidence for the psychometric soundness of the Artificial Intelligence Literacy Scale. Content validity was ensured through expert review during the item development process, allowing the scale items to reflect the conceptual scope of AI literacy as defined in the literature. Construct validity was supported by the convergence of exploratory and confirmatory factor analyses, which consistently identified the same five-factor structure across independent samples. Reliability analyses indicated high internal consistency for the overall scale and acceptable to high reliability coefficients for each subdimension. These findings are comparable to, and in some cases exceed, reliability values reported in existing AI literacy instruments, many of which focus on adult learners or preservice teachers. In addition, the high test–retest correlation suggests that the scale produces stable measurements over time, an aspect that is not consistently addressed in prior AI literacy research. Taken together, these results indicate that the developed instrument provides a reliable and valid means of assessing AI literacy among high school students, addressing a methodological gap identified in both national and international literature.

Group Differences in Artificial Intelligence Literacy

The group difference analyses revealed statistically significant variations in AI literacy scores across gender, grade level, school type, and patterns of AI use. These findings suggest that AI literacy among high school students is shaped by a combination of individual characteristics and contextual factors. Specifically, the finding that male students scored higher on Basic Knowledge and Awareness while female students scored higher on Ethical and Safe Use is broadly consistent with prior empirical work on AI literacy among secondary school students. A comparable pattern was reported by Zhong et al. (2025), who found that female secondary school students demonstrated lower AI literacy and AI career interest than their male peers, attributing this gap partly to socialized confidence differentials in STEM-related domains. Similarly, Ng et al. (2024) observed that boys demonstrated significantly higher AI knowledge scores, while girls scored higher on affective dimensions of AI literacy — a pattern that mirrors the gender differentiation observed across dimensions in the present study. These convergent findings suggest that gender differences in AI literacy are not uniform across all competency domains, and that educational interventions should be designed to address dimension-specific disparities rather than treating AI literacy as a monolithic construct. Importantly, the multidimensional structure of the scale allows for a more nuanced interpretation of these differences by identifying the specific dimensions in which variation occurs. This developmental pattern is consistent with findings reported by Zhong et al. (2025), who observed that upper secondary students demonstrated higher cognitive AI literacy skills compared to lower secondary students, suggesting that the capacity for critical and evaluative engagement with AI systems matures alongside broader cognitive development during adolescence. The present finding that grade-level differences were significant specifically for the Critical Perspective subdimension but not for more foundational dimensions is theoretically meaningful: it suggests that while basic AI knowledge may be accessible to students at various grade levels, the capacity to critically interrogate AI-generated content and recognize its epistemic limitations develops more gradually and may require sustained educational scaffolding. It should be noted, however, that this interpretation warrants caution given the unequal group sizes across grade levels, particularly the smaller number of 12th-grade participants ($n = 85$) relative to lower grade levels. This interpretation is supported by emerging evidence linking institutional AI readiness to student-level AI literacy outcomes. Recent multilevel analyses suggest that school-level AI capacity encompassing infrastructure, governance, and professional development plays a meaningful mediating role in shaping students' AI competencies, beyond individual-level factors alone (Guan et al., 2026). The present finding that Science High School students scored significantly higher than Anatolian and Vocational High School students on knowledge and application dimensions is consistent with this view: schools with stronger STEM orientations and more technologically enriched learning environments may provide students with greater opportunities for structured and critical engagement with AI tools. The post-hoc findings, particularly those related to Basic Knowledge and Awareness and Ethical and Safe Use, underscore the importance of examining subdimensions rather than relying solely on total scores. Notably, the finding that higher daily AI use was associated with lower Ethical and Safe Use scores — rather than higher — is particularly important from an educational standpoint. This inverse pattern suggests that increased exposure to AI tools does not automatically translate into more responsible or reflective use, a concern echoed in recent large-scale surveys of adolescent AI use (Pew Research Center, 2026). Students who use AI most frequently may develop procedural fluency without corresponding growth in ethical awareness, underscoring the need for instructional approaches that explicitly address ethical dimensions alongside skill development.

Methodological Contributions and Implications

Beyond substantive findings, this study makes an important methodological contribution by providing a psychometrically validated instrument tailored to the secondary education context. Unlike many existing measures that target adult populations or focus primarily on attitudes toward AI, the developed scale captures multiple dimensions of AI literacy relevant to high school students. By enabling the assessment of AI literacy at a granular level, the scale offers researchers and educators a tool for identifying specific areas of strength and need. For instance, a school counselor or curriculum coordinator could administer the scale at the beginning of an academic year to identify which AI literacy dimensions — such as ethical awareness or critical perspective — are least developed among students and then design targeted instructional modules or extracurricular activities to address these specific gaps. Similarly, a researcher evaluating the impact of an AI literacy intervention program could use the scale as both a pre- and post-measure, comparing subdimension scores before and after the intervention to assess which competency areas showed the greatest growth. This, in turn, can inform curriculum development, instructional design, and teacher education initiatives aimed at fostering ethical, critical, and applied AI competencies. From a policy perspective, the availability of a validated measurement tool supports evidence-based decision-making related to AI integration in secondary education. The scale's five-factor structurespanning Basic Knowledge and Awareness, Ethical and Safe Use, Application Skills, Critical Perspective, and Evaluation — provides a theoretically grounded and empirically validated framework that goes beyond existing instruments by capturing the full breadth of AI literacy competencies relevant to secondary school students. This granular measurement capability positions the scale as a uniquely valuable tool for both research and practice in AI literacy education.

Overall, the discussion demonstrates that artificial intelligence literacy among high school students emerges as a context-dependent and multidimensional competence shaped by individual engagement patterns, institutional environments, and opportunities for meaningful interaction with AI technologies.

Conclusion and Suggestions

This study aimed to develop a valid and reliable instrument to assess artificial intelligence literacy among high school students. In response to the growing integration of artificial intelligence technologies into educational and everyday contexts, a multidimensional scale was developed and psychometrically validated. The findings demonstrated that artificial intelligence literacy among secondary school students can be conceptualized through five interrelated dimensions: Basic Knowledge and Awareness, Ethical and Safe Use, Application Skills, Critical Perspective, and Evaluation. Exploratory and confirmatory factor analyses supported this structure, and reliability analyses indicated high internal consistency and temporal stability. Beyond its psychometric contributions, the study provides empirical evidence that artificial intelligence literacy varies across demographic and contextual variables, including gender, grade level, school type, and patterns of AI use. These findings underscore the importance of considering both individual and institutional factors when designing educational interventions aimed at fostering AI literacy in secondary education.

Limitations & Suggestions for Future Research

Several limitations should be considered when interpreting the findings of this study. First, the sample was limited to high school students in Türkiye, which may restrict the generalizability of the results to other national or cultural contexts. Although efforts were made to include participants from different school types, grade levels, and geographic regions, the sample remains subject to evidence heterogeneity that is, variability in students' educational experiences, technological access, and AI exposure across different school environments. This heterogeneity may introduce systematic differences that the current analytical framework cannot fully account for, and caution is therefore warranted when making broad generalizations from these findings. Second, data were collected using self-report measures, which are inherently susceptible to response bias, social desirability effects, and potential discrepancies between reported and actual AI literacy competencies. Students' self-assessments of their AI knowledge, ethical awareness, and evaluative skills may not fully correspond to their demonstrated abilities in authentic performance contexts. Third, the cross-sectional nature of the data limits causal interpretations regarding the development of AI literacy. While the findings reveal meaningful associations between AI literacy and demographic or usage-related variables, the study design does not permit conclusions about the direction or causal mechanisms underlying these relationships. Finally, the study focused on internal structure validity and internal consistency; future research may examine additional forms of validity and reliability, including criterion-related validity and cross-cultural applicability, using different samples and methods.

The findings of this study also suggest several directions for future research and practice. First, cross-cultural validation studies are needed to examine whether the five-factor structure of the Artificial Intelligence Literacy Scale holds across different national and educational contexts. Given that AI literacy is shaped by cultural norms, educational systems, and levels of technological infrastructure, the psychometric properties of the scale may vary meaningfully across settings, and international comparisons would both strengthen the instrument's global applicability and generate theoretically valuable insights.

Second, longitudinal research designs would allow researchers to track how AI literacy develops over time and to examine the role of formal instruction, informal AI exposure, and broader societal changes in shaping students' competencies. Whereas the present study provides a cross-sectional snapshot of AI literacy at a single point in time, longitudinal studies could illuminate developmental trajectories and identify critical periods for targeted educational intervention.

Third, the scale offers significant potential for use in intervention and experimental research. Researchers and practitioners could employ the instrument as a pre- and post-measure to evaluate the effectiveness of AI literacy curricula, teacher training programs, or school-wide digital literacy initiatives. Such research would not only generate evidence about what works in AI literacy education but also help educators identify specific subdimensions — such as ethical awareness or critical perspective — that may require more deliberate instructional attention.

Finally, from a policy standpoint, the availability of a validated and multidimensional AI literacy scale for secondary school students enables policymakers to monitor AI literacy levels at scale, identify at-risk groups, and design evidence-based strategies for integrating AI education into national curricula. The scale's subdimension structure makes it particularly well-suited for informing targeted policy decisions, as it allows for a granular understanding of where students' AI competencies are strongest and where gaps remain most pronounced.

In conclusion, the Artificial Intelligence Literacy Scale developed in this study provides a robust and context-appropriate measurement tool for secondary education. By enabling systematic assessment of AI literacy, the scale has the potential to support research, educational practice, and policy initiatives aimed at preparing students for responsible, critical and informed engagement with AI technologies not merely as passive participants in technology-driven environments, but as reflective agents capable of questioning, evaluating, and shaping the role of AI in their lives and communities (Pangrazio, 2026).

References

Baker, T., Smith, L., & Anissa, N. (2019). *Educ-AI-tion rebooted? Exploring the future of artificial intelligence in schools and colleges.* Nesta. https://www.nesta.org.uk/report/education-rebooted/

Bozkurt, A., & Sharma, R. C. (2024). Are we facing an algorithmic renaissance or apocalypse? Generative AI, ChatBots, and emerging human-machine interaction in the educational landscape. *Asian Journal of Distance Education, 19*(1), i–xii. https://doi.org/10.5281/zenodo.10791959

Byrne, B. M. (2016). *Structural equation modeling with AMOS: Basic concepts, applications, and programming* (3rd ed.). Routledge. https://doi.org/10.4324/9781315757421

Casal-Otero, L., Catala, A., Fernández-Morante, C., Taboada, M., Cebreiro, B., & Barro, S. (2023). AI literacy in K-12: A systematic literature review. *International Journal of STEM Education, 10*, Article 29. https://doi.org/10.1186/s40594-023-00418-7

DeVellis, R. F. (2016). *Scale development: Theory and applications* (4th ed.). SAGE Publications.

Elçiçek, M. (2024). Öğrencilerin Yapay Zeka Okuryazarlığı Üzerine Bir İnceleme. *Bilgi Ve İletişim Teknolojileri Dergisi, 6*(1), 24-35. https://doi.org/10.53694/bited.1460106

Eshet-Alkalai, Y. (2004). Digital literacy: A conceptual framework for survival skills in the digital era. *Journal of Educational Multimedia and Hypermedia, 13*(1), 93–106. https://www.learntechlib.org/primary/p/4793/

European Commission. (2022). *DigComp 2.2: The Digital Competence Framework for Citizens.* Publications Office of the European Union. https://doi.org/10.2760/115376

Fornell, C., & Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research, 18*(1), 39–50. https://doi.org/10.1177/002224378101800104

Guan, X., Zheng, M., Gašević, D., Guo, W., Liu, Y., Han, X., Gasevic, D., Ma, R., Wu, Q., & Yan, L. (2026). From school AI readiness to student AI literacy: A national multilevel mediation analysis of institutional capacity and teacher capability. *arXiv Preprint.* https://arxiv.org/abs/2603.20056

Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2019). *Multivariate data analysis* (8th ed.). Cengage Learning. https://www.cengage.uk/c/multivariate-data-analysis-8e-hair-babin-anderson-black/9781473756540/?searchIsbn=9781473756540

Hinkin, T. R. (1998). A brief tutorial on the development of measures for use in survey questionnaires. *Organizational Research Methods, 1*(1), 104–121. https://doi.org/10.1177/109442819800100106

Karaoğlan Yılmaz, F. G., & Yılmaz, R. (2023). Adaptation of Artificial Intelligence Literacy Scale into Turkish. *Journal of Information and Communication Technologies, 5*(2), 172-190. https://doi.org/10.53694/bited.1376831

Laupichler, M. C., Aster, A., Haverkamp, N., & Raupach, T. (2023). Development of the "scale for the assessment of non-experts' AI literacy" – An exploratory factor analysis. *Computers in Human Behavior Reports, 12*, Article 100338. https://doi.org/10.1016/j.chbr.2023.100338

Lintner, T. (2024). A systematic review of AI literacy scales. *npj Science of Learning, 9*, 50. https://doi.org/10.1038/s41539-024-00264-4

Long, D., & Magerko, B. (2020). What is AI literacy? Competencies and design considerations. *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems,* 1–16. https://doi.org/10.1145/3313831.3376727

Ng, D. T. K., Wu, W., Leung, J. K. L., Chiu, T. K. F., & Chu, S. K. W. (2024). Design and validation of the AI literacy questionnaire: The affective, behavioural, cognitive and ethical approach. *British Journal of Educational Technology, 55*(3), 1082–1104. https://doi.org/10.1111/bjet.13411

OECD. (2025). *Empowering learners for the age of AI: An AI literacy framework for primary and secondary education* (Review draft). OECD. https://ailiteracyframework.org

Pangrazio, L. (2026). The (im)possibility of AI literacy. *Learning, Media and Technology, 51*(1), 1–7. https://doi.org/10.1080/17439884.2026.2615553

Pew Research Center. (2026). *How teens use and view AI.* https://www.pewresearch.org/internet/2026/02/24/how-teens-use-and-view-ai/

Tabachnick, B. G., & Fidell, L. S. (2013). *Using multivariate statistics (6th ed.).* Pearson Education.

UNESCO. (2022). *K–12 AI curricula: A mapping of government-endorsed AI curricula.* United Nations Educational, Scientific and Cultural Organization. https://unesdoc.unesco.org/ark:/48223/pf0000380602

Wang, B., Rau, P. L. P., & Yuan, T. (2022). Measuring user competence in using artificial intelligence: Validity and reliability of artificial intelligence literacy scale. *Behaviour & Information Technology, 42*(9), 1324–1337. https://doi.org/10.1080/0144929X.2022.2072768

Wang, Z., Chai, C. S., Li, J., Leong, M. W. H., & Lim, C. P. (2025). Assessment of AI ethical reflection: The development and validation of the AI ethical reflection scale (AIERS) for university students. *International Journal of Educational Technology in Higher Education, 22*, Article 19. https://doi.org/10.1186/s41239-025-00519-z

Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education – Where are the educators? *International Journal of Educational Technology in Higher Education, 16*, Article 39. https://doi.org/10.1186/s41239-019-0171-0

Zhang, C., & Magerko, B. (2025). Generative AI literacy: A comprehensive framework for literacy and responsible use. *arXiv Preprint,* 9038. https://doi.org/10.48550/arXiv.2504.19038

Zhong, B., & Liu, X. (2025). Evaluating AI literacy of secondary students: Framework and scale development. *Computers & Education, 227,* 105230. https://doi.org/10.1016/j.compedu.2024.105230

Appendix 1. Artificial Intelligence Literacy Scale

Gender (Cinsiyet):

☐ Male (Erkek) ☐ Female (Kadın)

Grade Level (Sınıf Düzeyi):

☐ Preparatory Class ☐ 9th Grade (9. Sınıf) ☐ 10th Grade (10. Sınıf) ☐ 11th Grade (11. Sınıf) ☐ 12th Grade (12. Sınıf)

School Type (Okul Türü):

☐ Anatolian High School (Anadolu Lisesi) ☐ Science High School (Fen Lisesi) ☐ Vocational High School (Mesleki Lisesi) ☐ Other (Diğer)

Factor	Item	Item Statement
Factor 1: Basic Knowledge and Awareness (Temel Bilgi ve Farkındalık)	1	Yeni geliştirilen yapay zekâ uygulamalarını takip edebilirim. *(I can keep up with newly developed artificial intelligence applications.)*
	2	Yapay zekânın ne olduğunu açıklayabilirim. *(I can explain what artificial intelligence is.)*
	3	Yapay zekânın hangi alanlarda kullanıldığını açıklayabilirim. *(I can explain the areas in which artificial intelligence is used.)*
	4	Sosyal medya uygulamalarındaki yapay zekâ özelliklerini belirleyebilirim. *(I can identify artificial intelligence features in social media applications.)*
	5	Yapay zekâ araçlarının nasıl çalıştığını öğrenmek için araştırma yapabilirim. *(I can conduct research to learn how artificial intelligence tools work.)*
	6	Yapay zekâ kullanırken etkili sonuçlar elde etmek için ayrıntılı istemler (promptlar) yazabilirim. *(I can write detailed prompts to obtain effective results when using artificial intelligence.)*
	7	Yapay zekâ araçlarının kullanım talimatlarını kolayca anlayabilirim. *(I can easily understand the usage instructions of artificial intelligence tools.)*
Factor 2: Ethical and Safe Use (Etik ve Güvenli Kullanım)	8	Yapay zekâ kullanırken kişisel bilgilerimin gizliliğine dikkat ederim. *(I pay attention to the confidentiality of my personal information when using artificial intelligence.)*
	9	Yapay zekâ araçlarını kullanırken güvenliğime dikkat ederim. *(I pay attention to my security when using artificial intelligence tools.)*
	10	Yapay zekâyı kötü amaçlarla kullanmamaya özen gösteririm. *(I am careful not to use artificial intelligence for malicious purposes.)*
	11	Yapay zekâ araçlarını ödevlerimde kullanırken etik sınırları gözetmeye çalışırım. *(I try to observe ethical boundaries when using artificial intelligence tools for my assignments.)*
Factor 3: Application Skills (Uygulama Becerileri)	12	Ödevlerimi yaparken yapay zekâ destekli uygulamaları kullanabilirim. *(I can use artificial intelligence-supported applications while doing my assignments.)*
	13	Çevrim içi yapay zekâ araçlarını (örneğin ChatGPT, Google Gemini) günlük hayatımda kolayca kullanabilirim. *(I can easily use online artificial intelligence tools (e.g., ChatGPT, Google Gemini) in my daily life.)*
	14	Yapay zekâ araçlarını farklı derslerde (örneğin matematik veya dil öğrenimi) nasıl kullanacağımı bilirim. *(I know how to use artificial intelligence tools in different subjects (e.g., mathematics or language learning).)*
	15	Yapay zekâ araçlarını günlük yaşam problemlerini çözmek için kullanabilirim. *(I can use artificial intelligence tools to solve daily life problems.)*
Factor 4: Critical Perspective (Eleştirel Bakış Açısı)	16	Yapay zekâ araçlarını, halüsinasyon (gerçek dışı veya yanıltıcı bilgi üretimi) farkındalığıyla kullanabilirim. *(I can use artificial intelligence tools with an awareness of hallucinations (producing incorrect or misleading information).)*
	17	Yapay zekâ araçlarını, manipülasyon (yanıltıcı bilgi veya aldatıcı yönlendirme) farkındalığıyla kullanabilirim. *(I can use artificial intelligence tools with an awareness of manipulation (misleading information or deceptive guidance).)*
Factor 5: Evaluation (Değerlendirme)	18	Bir görselin yapay zekâ tarafından üretilip üretilmediğini anlayabilirim. *(I can recognize whether a visual has been generated by artificial intelligence.)*
	19	Yapay zekâ destekli araçlar tarafından üretilen çıktılardaki (ör. metin, görsel) hataları tespit edebilirim. *(I can detect errors in outputs (e.g., text, images) produced by artificial intelligence-supported tools.)*
	20	Yapay zekâ araçlarının her zaman doğru sonuçlar üretmediğini bilirim. *(I know that artificial intelligence tools do not always produce correct results.)*

Note. Items were rated on a 5-point Likert scale (1 = Strongly disagree / Kesinlikle katılmıyorum, 5 = Strongly agree / Kesinlikle katılıyorum).

About the Author(s):

Taha Aydogmus; Bahcesehir University, Turkey, taha.aydogmus@bahcesehir.edu.tr, https://orcid.org/0000-0002-0383-1999.

Yavuz Samur; Bahcesehir University, Turkey, yavuz.samur@bau.edu.tr, https://orcid.org/0000-0003-4269-7099.

Author's Contributions (CRediT)

Taha Aydogmuş contributed to the conceptualization, methodology design, investigation, data curation, analysis, visualization, and writing of the original draft, as well as the review and editing of the manuscript. Yavuz Samur provided supervision, proccess administration, methodological guidance, validation of analyses, and contributed to the review and editing process. All authors have read and agreed to the published version of the manuscript.

Sustainable Development Goals (SDGs)

This study is linked to the following SDG(s): Quality education (SDG 4) and Partnerships for the goals (SDG 17).

Authors' Disclosures

Editing and Review: This article was reviewed, edited, and improved with the help of a generative AI tool using ChatGPT 3.5 and ChatGPT o4-mini versions between February and April 2025 to complete the work of human editors in terms of grammar. Human authors critically evaluated and verified the content to maintain academic rigor. The authors also assessed and addressed potential biases inherent in AI-generated content. The final version of the article is entirely the responsibility of the human authors.

Data Accessibility Statement

The datasets used and/or analysed during the current study are available from the corresponding author on reasonable request.

Ethics and Consent

All procedures in this study were conducted in accordance with scientific research ethics. The data were used solely for academic purposes and remained fully confidential. Prior to data collection, participants were informed that their responses would be used exclusively for academic research, that participation was voluntary, and that all data would remain anonymous and confidential.

Funding Information

Not applicable.

Competing Interests

The authors have no competing interests to declare.

Article History

Received: January 12, 2026 – Accepted: May 6, 2026.

Suggested citation:
Aydogmus, T., & Samur, Y. (2026). AI Literacy Level of Students: A Scale Development Study. *Asian Journal of Distance Education, 21*(1), 171-204. https://doi.org/10.5281/zenodo.20045416