

Article

Validation of the Meta AI Literacy Scale (MAILS) in Ecuadorian High School Students: Evidence of Validity, Measurement Invariance, and Latent Profiles

Ángel-Freddy Rodríguez-Torres ¹, Diana Carolina Treviño-Villarreal ², Karla Paulina Hidalgo-Montesinos ³
and José-Antonio Marín-Marín ^{4,*}

¹ Faculty of Physical Culture, Central University of Ecuador, Quito 170129, Ecuador; afrodriguez@uce.edu.ec

² Directorate of Fund Management, Autonomous University of Nuevo León, Monterrey 66455, Mexico; diana.trevinov@uanl.mx

³ Faculty of Philosophy, Letters and Educational Sciences, Central University of Ecuador, Quito 170129, Ecuador; kphidalgo@uce.edu.ec

⁴ Department of Didactics and School Organisation, Faculty of Education Sciences, University of Granada, 18071 Granada, Spain

* Correspondence: jmarin@ugr.es

Abstract

This study aimed to validate the Meta AI Literacy Scale (MAILS) in Ecuadorian high school students and examine latent profiles of Artificial Intelligence (AI) literacy. An instrumental, cross-sectional study was conducted with a sample of 971 students aged 14 to 19 years. Confirmatory factor analysis (CFA) supported a nine-factor correlated model with satisfactory fit indices (CFI = 0.953, TLI = 0.946, RMSEA = 0.053, SRMR = 0.058), high internal consistency (total α = 0.973; ω = 0.974), and scalar measurement invariance by sex and type of educational centre. Latent profile analysis (LPA) identified four differentiated profiles organised along a quantitative continuum of competence. Overall, the findings provide empirical evidence regarding the validity and utility of the MAILS in a sample of Ecuadorian high school students, providing a basis for future validations in other Spanish-speaking and Latin American contexts, and suggest that intentional academic use of the internet is associated with higher levels of AI literacy, although with small effect sizes. The validated instrument constitutes a resource for educational researchers and practitioners seeking to characterise AI literacy in secondary school populations and to inform evidence-based curricular interventions.

Keywords: digital literacy; artificial intelligence; high school; psychometric validation; latent profile analysis



Academic Editors: Kristen Gregory,
Xi Lin and Ken Luterbach

Received: 11 May 2026

Revised: 22 June 2026

Accepted: 27 June 2026

Published: 2 July 2026

Copyright: © 2026 by the authors.
Licensee MDPI, Basel, Switzerland.
This article is an open access article
distributed under the terms and
conditions of the [Creative Commons
Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

1. Introduction

Over recent decades, a transition towards the digital era has taken place, driven by the convergence of technological advances, globalisation, and increasing competitiveness, as well as by the widespread adoption of Information and Communication Technologies (ICT). These are considered indispensable tools for performing multiple tasks simultaneously, searching and managing information, developing digital educational activities, and supporting commerce, work, and communication processes, among other domains. Within this context of technological transformation, artificial intelligence (AI), together with ICT, has gained growing relevance (Cruz-Pérez et al., 2019; Lampropoulos et al., 2026; Licona-Quiterio, 2024). This rapid expansion raises a central question for educational

research: how can students' capacity to use, understand, and critically engage with AI be systematically assessed? Addressing this question requires validated instruments calibrated for the specific educational and cultural contexts in which AI literacy is being developed.

In this context, AI may be defined as the capacity of machines to carry out tasks that typically require human cognitive skills. It is a broad concept that integrates diverse algorithms and techniques aimed at simulating intelligent behaviour. These systems do not possess consciousness or understanding in the human sense, but operate through data processing and mathematical models, and hence have no autonomous will (López de Mántaras & Brunet-Crosa, 2023; Suárez-Muñoz & Bautista-Tejeda, 2024).

This capacity facilitates the development of valuable services and is embedded in various aspects of daily life, enabling machines to simulate complex processes related to problem-solving, as well as tasks in creative and linguistic domains (UNESCO, n.d.-a). These systems operate without autonomous will or genuine understanding, but their growing embedding across services, creative domains, and problem-solving tasks makes it essential to equip students with the competencies needed to use them effectively and critically.

From a technical perspective, many new AI systems are grounded in the use of artificial neural networks, defined as computational models inspired by the functioning of the biological nervous system. These networks analyse relationships between inputs and outputs within large volumes of data, facilitating learning processes based on pattern identification used for problem-solving (Guerrero-González & Tendilla-Beltrán, 2022; López de Mántaras & Brunet-Crosa, 2023).

AI has transformed different sectors, one of which is education (Vieriu & Petrea, 2025). In this domain, the use of AI-powered technological tools supports the personalisation of teaching and learning processes by adapting to students' specific needs and promoting greater engagement (Tamares-Peña & García-López, 2024). Moreover, this technology redefines the ways in which students access knowledge, interact with academic content, and experience assessment processes (Mora et al., 2024). This rapid integration of AI tools in educational settings highlights the urgency of developing and validating instruments capable of assessing students' AI literacy competencies in a reliable and contextually sensitive manner.

1.1. AI in the Educational Field

From an educational perspective, AI presents notable potential for addressing major challenges and driving innovation across various pedagogical practices (UNESCO, n.d.-b). It is a significant tool for improving education and preparing students for the future, offering personalised and adaptive feedback and enhancing administrative efficiency (Bolaño-García & Duarte-Acosta, 2024; Paredes-Gallardo, 2024; Pozo-Sánchez et al., 2026). Nevertheless, its application varies across countries, reflecting disparities related to technological infrastructure, digital instruction, and funding, among other factors (UNESCO, 2024), and considerable expectations have been generated regarding its future impact (Holmes & Tuomi, 2022).

It is worth noting that AI in education also intervenes in the cultural context and in cognitive processes and is thus consolidating as an essential research focus in educational computing (Hwang et al., 2020).

Authors such as Bolaño-García and Duarte-Acosta (2024) have argued that it is a priority to adopt a critical and balanced perspective during the implementation of AI in schools, to ensure that both teachers and students understand its limitations and benefits and can use it effectively in the teaching–learning process.

Paredes-Gallardo (2024), Suárez-Estavillo (2025), and Tamares-Peña and García-López (2024) highlight the existence of various barriers related to technical understanding, ethical and cultural acceptance, data bias, student privacy, resistance to change, teacher training,

and accessibility to resources. All of this occurs despite the significant potential of AI, underscoring the importance of an appropriate approach and adequate teacher preparation.

Vieriu and Petrea (2025) argue that AI offers personalised learning, improved academic performance, and greater student participation. However, challenges are also identified, including potential excessive dependency, reduced critical thinking, academic dishonesty, and risks associated with data privacy. This highlights the importance of creating a well-structured framework, supported by ethical guidelines, to maximise benefits and mitigate risks.

1.2. AI in Secondary Education

The integration of AI into secondary education systems represents an unprecedented paradigm shift, in which digital tools have ceased to be mere supports and become active cognitive agents (OECD, 2026). During 2025–2026, a rapid expansion of AI use in education has been documented. OECD (2026) reports that, according to a Vodafone Foundation (2025) survey of over 7000 students aged 12–17 across seven European countries, 48% declared having used ChatGPT in 2024; among teachers, OECD's TALIS 2024 survey found that 37% of lower secondary teachers were already using AI tools in their work. This rapid adoption has generated concern about what has been termed the “performance trap”: a critical gap between immediate task performance and genuine long-term learning (OECD, 2025a).

A field experiment conducted in Türkiye and cited by the OECD (2026) found that access to GPT-4 improved short-term task performance by 48% with a standard interface, yet students performed 17% worse in assessments conducted without AI assistance once access was removed; this evidence suggests that AI use without explicit pedagogical design can undermine genuine learning, a pattern the OECD (2026) terms “metacognitive laziness”. This phenomenon, associated with “metacognitive laziness”, occurs when AI is used as a shortcut to obtain final answers rather than as a partner in the cognitive process (Fuster Rabella, 2025). In response to this risk of intellectual atrophy, literacy frameworks such as the AILit Framework propose that secondary education must prioritise the development of durable skills and critical competencies over mere technical operability (OECD, 2025a).

A further concern is the emerging notion of a “second digital divide”, in which inequality may reside not in access but in the quality and purpose of use, with some evidence suggesting that more advantaged groups tend to employ AI for cognitive augmentation while others lean towards passive consumption (OECD, 2026); this interpretive hypothesis, however, requires further empirical verification. The role of teachers, therefore, evolves as they become responsible for designing process-oriented assessments that protect human agency and foster deep thinking (OECD, 2026; Schwartz & Diliberti, 2026). Sixty per cent of students already express concern that AI may harm their critical thinking skills, underscoring the urgency of ethical pedagogical mediation (Schwartz & Diliberti, 2026).

In this regard, Zhong and Liu (2025) carried out research demonstrating that secondary school students show a high disposition to be educated in AI literacy. They also highlighted the need to develop instruments adjusted to the age, cognitive stages, and experiences of students at this educational level. Along similar lines, the reality of many secondary schools located in low-income contexts evidences a lack of resources and skills to guide students in the current digital environment. Consequently, if this tendency persists, AI could exacerbate existing inequalities (Dúo Terrón et al., 2023; Goth-Itoi, 2023).

1.3. AI Literacy

AI literacy refers to the set of basic competencies that are essential for people to know, understand, use, apply, evaluate, and create AI systems. These competencies integrate elements of data science, computational thinking, and a multidisciplinary approach (Ng

[et al., 2021](#)), and also include the ability to critically analyse AI's impact in ethical, social, and cultural domains ([Buitrago et al., 2024](#)).

In this sense, society requires cognitive training that allows individuals to face technological challenges in their respective professional fields, as well as to promote responsibility and ethical awareness in the use of AI ([Ng et al., 2021](#)).

Several instruments have been proposed to measure AI literacy in educational contexts, each with a specific scope and psychometric characteristics. [Ng et al. \(2021\)](#) developed a theoretical framework articulating four competency dimensions (Know & Understand AI, Use & Apply AI, Evaluate & Create AI, and AI Ethics), subsequently operationalised into measurement scales for young adults in Hong Kong. [Carolus et al. \(2023\)](#) proposed the MAILS, validated with German university students ($n = 1000$, $M = 24.6$ years), covering nine dimensions across 34 items on an 11-point scale, with evidence of convergent and discriminant validity. [Koch et al. \(2024\)](#) developed an abbreviated 15-item version of the MAILS with improved parsimony. [Chiu et al. \(2024\)](#) proposed an AI literacy scale for secondary students in Taiwan, while [Zhong and Liu \(2025\)](#) validated a similar instrument with Chinese adolescents. Notably, none of these validated instruments has been adapted and validated in Spanish-speaking or Latin American secondary school contexts, nor with students from highly heterogeneous institutional settings such as those of Ecuador, which constitutes the primary justification for the present study.

Despite growing research on AI literacy, studies that validate multidimensional instruments in Spanish-speaking school populations and Latin American contexts remain scarce. This gap is especially relevant in secondary education, where students begin to use AI tools for academic purposes before having consolidated critical literacy ([OECD, 2025b](#); [Zhong & Liu, 2025](#)). In response to this need, the present study aims to provide psychometric evidence on the validity, reliability, and invariance of the Meta AI Literacy Scale (MAILS) in Ecuadorian high school students, as well as to characterise latent profiles of AI literacy. This contribution is pertinent both for educational research and for curricular design in secondary education, a domain in which AI literacy is beginning to consolidate as a formative priority at the international level ([Chiu et al., 2024](#); [Ng et al., 2021](#); [OECD, 2025b](#)).

The general objective of this study is to validate the factorial structure, reliability, and invariance of the Meta AI Literacy Scale (MAILS) in high school students, as well as to characterise metacognitive AI literacy profiles based on their dimensions and academic use patterns. The following specific objectives are derived from this general objective:

SO1: To analyse the factorial structure and reliability of the MAILS in high school students.

SO2: To examine measurement invariance by sex and type of educational centre, where applicable.

SO3: To identify student profiles based on their scores on the MAILS dimensions.

SO4: To examine differences in total scores (MAILS_TOTAL) and subscales according to contextual variables and AI use patterns.

2. Materials and Methods

This study adopts an instrumental design, oriented towards the psychometric validation of the Meta AI Literacy Scale (MAILS; [Carolus et al., 2023](#)) in a high school population, with the aim of obtaining evidence of validity, reliability, and factorial invariance in the Ecuadorian educational context. This study is also cross-sectional in nature, as data were collected at a single time point, and employs a quantitative, psychometric-inferential analytical approach.

2.1. Participants

The sample consisted of 971 high school students from Ecuador, selected through non-probabilistic convenience sampling based on criteria of accessibility and availability. Students from state-funded (fiscal), municipal, and private institutions participated. Inclusion criteria comprised enrolment in Years 1, 2, or 3 of high school, informed consent from families, and complete questionnaire responses. About ethical procedures, the study was conducted in accordance with the Declaration of Helsinki and approved by the relevant Ethics Committee (protocol code 001-PSN-2024). Given that the sample included minor participants, written informed consent was obtained from all legal guardians or parents prior to data collection. Additionally, student assent was obtained from all minor participants, who were informed of the voluntary nature of their participation, their right to withdraw at any time without penalty, and the confidential and anonymous treatment of their data. Adult participants (aged 18 or over) provided their own informed consent directly. It should be noted that this sampling strategy does not allow generalisation of results to the Ecuadorian education system as a whole; findings are valid for the analysed sample and should be treated as exploratory evidence requiring replication in more representative samples.

Data collection was carried out in October 2025. Subsequently, a database depuration process was performed, verifying the consistency and integrity of responses. Of a total of 980 questionnaires collected, 971 were considered valid; records with substantive omissions or inconsistencies were excluded.

Regarding sex distribution, 45.7% ($n = 444$) of participants were female, and 54.3% ($n = 527$) were male. Participants' ages ranged from 14 to 19 years, with a mean age of 16.5 years ($SD = 1.00$), indicating moderate dispersion.

With respect to the type of educational centre, 80.3% ($n = 780$) of students attended state-funded institutions, 11.5% ($n = 112$) attended municipal centres, and 8.1% ($n = 79$) attended private institutions. This distribution reflects a marked imbalance across groups, particularly for the municipal and private categories, which should be borne in mind when interpreting measurement invariance results by institution type: the smaller group sizes reduce statistical power for detecting non-invariance and may affect the robustness of group comparisons. Regarding school year, 20.0% ($n = 194$) were in Year 1 of high school, 19.0% ($n = 185$) in Year 2, and 61.0% ($n = 592$) in Year 3. Furthermore, 42.0% ($n = 408$) of participants were enrolled in the Science track, 54.6% ($n = 530$) in the Technical track, and 3.4% ($n = 33$) in the International Baccalaureate track.

2.2. Instrument

The Meta AI Literacy Scale (MAILS), developed by Carolus et al. (2023), was used. This is a self-report instrument designed to evaluate AI literacy from a multidimensional perspective, integrating technical, ethical, and psychological components. The choice of MAILS is grounded in its conceptual breadth and in the prior evidence of validity reported in recent studies (Carolus et al., 2023; Koch et al., 2024).

From a structural perspective, the scale assesses AI literacy through 9 dimensions (7 primary dimensions and 2 composed of 2 subdimensions each): Use and Application, Understanding, Detection, Ethics, Creation, AI Self-Efficacy (Problem Solving, AI Self-Efficacy), AI-Mediated Learning, Own Competence (AI Persuasion Literacy and Own Competence) and Emotional Regulation in AI Interaction. The instrument comprises 34 items covering these dimensions.

Items use an 11-point Likert response format, ranging from 0 = "never or very rarely" to 10 = "very often". This format enables differentiated scores for each dimension, as well as a global AI literacy score, facilitating a detailed analysis of the construct.

In the present study, the scale was adapted linguistically and contextually to the Ecuadorian educational context while respecting the original instrument's structure. The adaptation followed international recommendations for the translation and adaptation of instruments (International Test Commission, 2017; Beaton et al., 2000), comprising four sequential phases.

First, two independent forward translations of the 34 original English items into Spanish were produced by two bilingual translators with expertise in educational technology and applied linguistics, and a consensus version was agreed upon by the research team.

Second, the consensus version was reviewed by a panel of seven experts (specialists in educational technology and secondary education), with a mean of 16.3 years of professional experience, who evaluated each item for semantic, idiomatic, and conceptual equivalence with the original English version, as well as for its relevance and clarity for the Ecuadorian secondary school population. Experts rated each item using a 4-point relevance/clarity scale (1 = not relevant/clear, 4 = highly relevant/clear), and items with an inter-rater agreement below 0.80 were revised or reformulated.

Third, a semantic clarity assessment (pilot test) was conducted with 31 students with characteristics similar to those of the final sample (age range, type of institution), who completed the adapted instrument and reported any item perceived as ambiguous, difficult to understand, or culturally inadequate. Based on this pilot, four items (MAILS_U4, MAILS_C2, MAILS_C4 and MAILS_D2) were minimally reworded from a linguistic standpoint to improve comprehensibility, mainly affecting the Understanding, Creativity and Detection dimensions.

Fourth, the resulting Spanish (Ecuadorian) version, comprising the same 34 items and nine-dimension structure as the original scale, was adopted for data collection. The full adapted version, with item wording, dimension membership, and corresponding variable codes, is provided in Appendix A.

Regarding psychometric properties, prior studies have evidenced that the MAILS presents adequate levels of validity and reliability, supported by factor analyses and satisfactory internal consistency values across its dimensions. Reliability estimates for the present sample (Cronbach's alpha and McDonald's omega; Tables 1 and 2) are reported here—rather than exclusively in the Results section—to contextualise the instrument's suitability prior to CFA, following the convention of reporting preliminary psychometric properties as part of instrument description in validation studies (Hernández-Nieto, 2002). These values should nonetheless be interpreted alongside the full set of validity evidence reported in the Results. It is noted that α and ω values above 0.86 across all dimensions, while indicating strong internal consistency, may also reflect some degree of item intercorrelation that warrants caution against interpreting reliability alone as evidence of construct validity; item elimination analysis confirmed, however, that no individual item substantially reduced reliability, suggesting that the observed values reflect genuine construct coherence rather than problematic item redundancy.

Table 1. Reliability statistics for the MAILS.

Scale	Cronbach's α	McDonald's ω
MAILS (total scale)	0.973	0.974

Note. α = Cronbach's alpha; ω = McDonald's omega.

Table 2. Internal consistency of the MAILS dimensions.

Dimension/Subdimension	No. Items	α	ω
Use and Application of AI (A)	6	0.939	0.939
Understanding of AI (U)	6	0.928	0.928
Detection of AI Systems (D)	3	0.862	0.864
AI Ethics (E)	3	0.901	0.902
Creation of AI (C)	4	0.926	0.928
AI Self-Efficacy—Problem Solving (PS)	3	0.902	0.903
AI Self-Efficacy—AI-Mediated Learning (L)	3	0.915	0.915
Own Competence—AI Persuasion Literacy (P)	3	0.872	0.873
Own Competence—Emotional Regulation (R)	3	0.907	0.908

Note. α = Cronbach's alpha; ω = McDonald's omega.

2.3. Data Analysis

Statistical analyses were conducted following a sequential approach oriented towards psychometric validation of the instrument, evaluation of its stability across groups, and analysis of latent profiles associated with AI literacy.

First, descriptive analyses were carried out to examine the distribution of variables, as well as internal consistency of the scale using Cronbach's alpha (α) and McDonald's omega (ω), both at the item and dimension level.

Subsequently, construct validity was evaluated through a confirmatory factor analysis (CFA), specifying a model of nine correlated latent factors in accordance with the original theoretical structure of the instrument. Model fit was examined using absolute and incremental goodness-of-fit indices, including chi-square (χ^2), the Comparative Fit Index (CFI), the Tucker–Lewis Index (TLI), the Root Mean Square Error of Approximation (RMSEA) with its 90% confidence interval, and the Standardised Root Mean Square Residual (SRMR).

Convergent validity was assessed through the Average Variance Extracted (AVE), with values equal to or above 0.50 considered indicative of adequacy. Discriminant validity was examined through the heterotrait–monotrait ratio (HTMT), using values below 0.85 as a conservative criterion and values below 0.90 as a more liberal criterion.

A multigroup factorial invariance analysis was then conducted to evaluate the equivalence of the measurement model across groups (sex and type of educational centre). Configural, metric, and scalar models were contrasted hierarchically, with a change in CFI ($\Delta\text{CFI} \leq 0.01$) as the primary criterion for accepting invariance between successive models.

Subsequently, a Latent Profile Analysis (LPA) was applied to identify homogeneous patterns of AI literacy in the sample. Optimal profile number selection was based on information criteria (AIC, BIC, and SABIC), entropy, and relative class sizes.

Finally, to test hypotheses regarding differences between profiles, one-way analyses of variance (ANOVA) were performed, together with the calculation of effect size using eta squared (η^2). A multinomial logistic regression was also estimated to examine the predictive value of academic internet use and AI tool diversity on profile membership, reporting logit coefficients (β), standard errors, z-statistics, p-values, odds ratios (OR), and their 95% confidence intervals.

Interpretation criteria followed widely accepted recommendations in structural modelling and psychometric research (Chen, 2007; Cheung & Rensvold, 2002; Hu & Bentler, 1999; Kline, 2023).

Descriptive analyses, internal consistency estimation (Cronbach's alpha and McDonald's omega), CFAs, convergent and discriminant validity evaluation (including HTMT), and multigroup factorial invariance analyses were conducted with JASP (version 0.96) (JASP Team, 2025). One-way ANOVAs and multinomial logistic regression were also performed with JASP. Latent profile analysis was conducted in R (version 4.5.2) (R Core Team,

2024), using RStudio (version 2026.01.0+392, Apple Blossom) as a development environment. Latent profile models were estimated using the tidyLPA package (Rosenberg et al., 2018), supported by the mclust package (Scrucca et al., 2016) for finite mixture modelling. Data preparation and management were performed with the tidyverse ecosystem, and graphical representation of latent profiles was produced using ggplot2 (Wickham, 2016).

All analyses were conducted at a statistical significance level of $p = 0.05$.

3. Results

To characterise the study participants, several questions were included to identify the habitual device used to access the internet and to quantify daily connection hours dedicated to leisure, academic, and research purposes. The majority of participants access the internet via mobile phone (82.4%, $n = 800$), followed by computer (17.0%, $n = 165$), with only six participants using a tablet (0.6%). Regarding internet connection hours and use patterns, Table 3 shows that 35.3% of participants dedicate 2–4 h to leisure purposes, with 10.1% spending more than 8 h online for leisure-related activities. Regarding academic use, 36.4% also dedicate 2–4 h, while for research, figures are lower, with 37.8% spending less than 2 h.

Table 3. Frequency distribution and percentages of internet use by purpose ($n = 971$).

Hours Interval	Leisure n (%)	Academic n (%)	Research n (%)
<2 h	203 (20.9)	197 (20.3)	367 (37.8)
2–<4 h	343 (35.3)	353 (36.4)	332 (34.2)
4–<6 h	252 (26.0)	272 (28.0)	159 (16.4)
6–<8 h	75 (7.7)	97 (10.0)	73 (7.5)
≥ 8 h	98 (10.1)	52 (5.4)	40 (4.1)
Total	971 (100)	971 (100)	971 (100)

Note. Data are presented as absolute frequencies (n) and percentages (%). Intervals correspond to daily hours of internet use by purpose (leisure, academic, research). Percentages were calculated over the total sample.

To contextualise the technological profile of the sample, patterns of AI tool use and the main academic purposes associated with their use were analysed (see Table 4). As both questions allowed multiple selection, case percentages may exceed 100%.

Table 4. Patterns of AI tool use and purposes ($n = 971$).

Category	n	% Cases
Type of AI tool		
Conversational assistants	877	92.1
Translation tools	291	30.6
Information analysis/search	256	26.9
Writing and text editing	159	16.7
Audiovisual generation	69	7.2
SEO/plagiarism detection	42	4.4
Reasons for AI use		
Research	783	80.6
Completing assignments	668	68.8
Searching for information	483	49.7
Planning	171	17.6
Improving writing	69	7.1
Other academic purposes	—	<6% each

Note. Multiple-response questions; percentages may exceed 100%.

Results show a clear predominance of conversational assistants, used by over 90% of participants. Regarding purposes, AI use is oriented primarily towards research tasks,

completing assignments, and searching for information. These patterns confirm a predominantly academic and functional use of AI, as reported through self-report measures; actual patterns of use may differ from self-perceived ones and should be corroborated through observational or log-based data in future studies, which is consistent with the subsequent analysis of AI literacy profiles.

3.1. Confirmatory Factor Analysis

Following verification of the instrument's internal consistency, construct validity was evaluated through a CFA to empirically test the adequacy of the proposed theoretical structure for the MAILS in the analysed sample.

A nine-factor correlated CFA model was estimated, in accordance with the original instrument structure. Specifically, the model included: (a) Use and Application of AI, (b) Understanding of AI, (c) Detection of AI Systems, (d) AI Ethics, (e) Creation of AI, (f) AI Self-Efficacy—Problem Solving, (g) AI Self-Efficacy—AI-Mediated Learning, (h) Own Competence—AI Persuasion Literacy, and (i) Own Competence—Emotional Regulation. Each item was specified as an exclusive indicator of its corresponding theoretical factor, with correlations permitted among all latent factors.

The model was estimated using a robust Maximum Likelihood estimator (MLM), which provides standard errors and a mean-adjusted test statistic robust to non-normality. This estimator was selected, instead of conventional ML, given that the items were measured on an 11-point Likert-type scale; although treated as continuous given the large number of response categories and the sample size ($n = 971$), robust estimation safeguards inferential accuracy against potential departures from multivariate normality. Univariate skewness and kurtosis were examined for all 34 items prior to estimation (Appendix B, Table A2), ranging from -0.41 to 0.35 and from -1.14 to -0.70 , respectively—well within the conservative thresholds proposed for ML-based estimation (skewness < 3 , kurtosis < 10 ; Kline, 2023), indicating approximate univariate normality and supporting the adequacy of the estimation procedure. Model identification was ensured by fixing the variance of each latent factor to 1, enabling factor standardisation and facilitating interpretation of factor loadings and inter-factor covariances (Clogg & Bollen, 1991).

The exact model fit test was statistically significant, $\chi^2(491) = 1957, p < 0.001$. This result was expected given the sensitivity of the χ^2 statistic to sample size and model complexity. Model fit was therefore evaluated primarily from complementary fit indices.

Incremental and absolute fit indices indicated good to excellent model fit: CFI = 0.953, TLI = 0.946, RMSEA = 0.053 (90% CI [0.038, 0.055]), and SRMR = 0.058 (see Table 5). These values fall within the criteria recommended in the specialised literature, supporting the adequacy of the proposed factorial model. AIC = 129,096 and BIC = 129,769 are also reported as references for comparison with alternative models.

All factor loadings were positive and statistically significant ($p < 0.001$), with high Z-statistics. Standardised loadings ranged from 0.75 to 0.93, with item-level R^2 ranging from 0.57 to 0.86, indicating that between 57% and 86% of each item's variance was explained by its corresponding latent factor. The complete table of standardised loadings, standard errors, Z values, and R^2 for all 34 items is provided in Appendix C (Table A3). No items with weak or problematic loadings were identified, providing evidence in favour of the instrument's convergent validity at the factorial level.

Convergent validity was evaluated using the Average Variance Extracted (AVE). As shown in Table 6, all MAILS dimensions present AVE values above the recommended threshold of 0.50, ranging from 0.678 to 0.782. These results indicate that items explain a substantial proportion of the variance of their respective latent constructs, supporting adequate convergent validity of the model (Fornell & Larcker, 1981; Hair et al., 2019).

Table 5. Goodness-of-fit indices for the confirmatory factor model.

Index	Obtained Value	Recommended Criterion	Reference
χ^2	1957	—	—
df	491	—	—
χ^2/df	3.985	<5 (acceptable)	Kline (2023)
CFI	0.953	≥ 0.90 (adequate); ≥ 0.95 (excellent)	Hu and Bentler (1999)
TLI	0.946	≥ 0.90 (adequate); ≥ 0.95 (very good)	Hu and Bentler (1999)
RMSEA	0.053	≤ 0.08 (adequate); ≤ 0.06 (good)	Browne and Cudeck (1993); Hu and Bentler (1999)
90% CI RMSEA	[0.038, 0.055]	Upper limit < 0.08	MacCallum et al. (1996)
SRMR	0.058	≤ 0.08 (adequate); ≤ 0.06 (good)	Hu and Bentler (1999)
AIC	129,096	Lower = better fit	Shipley and Douma (2021)
BIC	129,769	Lower = better fit	Shipley and Douma (2021)

Note. χ^2 = chi-square; df = degrees of freedom; CFI = Comparative Fit Index; TLI = Tucker–Lewis Index; RMSEA = Root Mean Square Error of Approximation; SRMR = Standardised Root Mean Square Residual. Fit criteria are based on widely accepted recommendations in the structural equation modelling literature.

Table 6. Convergent validity of the MAILS (Average Variance Extracted).

Dimension/Factor	AVE
Use and Application of AI (Apply AI—A)	0.721
Understanding of AI (Understand AI—U)	0.684
Detection of AI Systems (Detect AI—D)	0.678
AI Ethics (AI Ethics—E)	0.755
Creation of AI (Create AI—C)	0.768
AI Self-Efficacy—Problem Solving (AI Problem Solving—PS)	0.756
AI Self-Efficacy—AI-Mediated Learning (Learning—L)	0.782
Own Competence—AI Persuasion Literacy (AI Persuasion Literacy—P)	0.701
Own Competence—Emotional Regulation (AI Emotion Regulation—R)	0.766

Note. AVE = Average Variance Extracted. Values ≥ 0.50 indicate adequate convergent validity (Fornell & Larcker, 1981).

Discriminant validity was evaluated using the HTMT (Henseler et al., 2015). All values fell below the liberal threshold of 0.90 (Henseler et al., 2015), and the majority also fell below the more conservative criterion of 0.85 (Hair et al., 2019). Two values approached 0.89, near the upper limit of the liberal threshold, indicating that discriminant validity between those factors, while acceptable, is not unequivocal; this proximity is acknowledged as a limitation (see Discussion). None exceeded the 0.90 cut-off points established in the literature (see Table 7). Overall, these results support adequate empirical differentiation among the model’s constructs.

Table 7. Discriminant validity of the MAILS using the HTMT index.

Factor	F1	F2	F3	F4	F5	F6	F7	F8	F9
F1	—								
F2	0.825	—							
F3	0.715	0.862	—						
F4	0.758	0.889	0.889	—					
F5	0.500	0.544	0.535	0.539	—				
F6	0.768	0.850	0.844	0.845	0.589	—			
F7	0.664	0.733	0.741	0.748	0.677	0.807	—		
F8	0.510	0.657	0.685	0.710	0.320	0.661	0.529	—	
F9	0.585	0.685	0.696	0.726	0.392	0.742	0.618	0.770	—

Note. HTMT = heterotrait–monotrait ratio. Values below 0.85 indicate adequate discriminant validity (Henseler et al., 2015).

3.2. Measurement Invariance Analysis

To evaluate the equivalence of factorial structure across groups (SO2), multigroup factorial invariance analyses were conducted following a progressive hierarchical procedure, contrasting configural, metric, and scalar models successively.

For sex invariance, the configural model showed adequate fit, $\chi^2(982) = 2782.18$, CFI = 0.942, TLI = 0.934, RMSEA = 0.061 (90% CI [0.059, 0.064]), and SRMR = 0.042, indicating that the pattern of item–factor relationships is equivalent between males and females (see Table 8).

Table 8. Factorial invariance of the MAILS by sex.

Model	χ^2	df	CFI	TLI	RMSEA	90% CI RMSEA	SRMR	Δ CFI
Configural	2782.18	982	0.942	0.934	0.061	[0.059, 0.064]	0.042	—
Metric	2796.65	1007	0.943	0.936	0.061	[0.058, 0.063]	0.042	0.001
Scalar	2839.78	1032	0.942	0.937	0.060	[0.057, 0.063]	0.042	0.001

Note. Δ CFI calculated relative to the previous model. Invariance was accepted when CFI decrease between consecutive models did not exceed 0.01 (Cheung & Rensvold, 2002; Chen, 2007).

When equality constraints were imposed on factor loadings (metric model), fit remained practically unchanged ($\chi^2(1007) = 2796.65$, CFI = 0.943, RMSEA = 0.061), with a minimal change in CFI (Δ CFI = 0.001), clearly below the recommended threshold of 0.010 (Cheung & Rensvold, 2002; Chen, 2007).

Similarly, the scalar model, which added constraints on intercepts, showed equally adequate fit ($\chi^2(1032) = 2839.78$, CFI = 0.942, RMSEA = 0.060), with a trivial CFI change (Δ CFI = 0.001). These results support configural, metric, and scalar invariance by sex.

For invariance by type of educational centre, the configural model showed acceptable fit, $\chi^2(1473) = 3591.22$, CFI = 0.933, TLI = 0.924, RMSEA = 0.067 (90% CI [0.064, 0.069]), and SRMR = 0.043 (see Table 9).

Table 9. Factorial invariance of the MAILS by type of educational centre.

Model	χ^2	df	CFI	TLI	RMSEA	90% CI RMSEA	SRMR	Δ CFI
Configural	3591.22	1473	0.933	0.924	0.067	[0.064, 0.069]	0.043	—
Metric	3635.51	1523	0.933	0.926	0.065	[0.063, 0.068]	0.044	0.000
Scalar	3692.13	1573	0.933	0.929	0.065	[0.062, 0.067]	0.044	0.000

Note. Δ CFI calculated relative to the previous model. Invariance was accepted when CFI decrease between consecutive models did not exceed 0.01 (Cheung & Rensvold, 2002; Chen, 2007).

Imposing equality constraints on factor loadings (metric model) did not produce substantial variations in fit (CFI = 0.933; Δ CFI = 0.000), evidencing metric invariance. Similarly, the scalar model maintained equivalent fit levels (CFI = 0.933; Δ CFI = 0.000), supporting scalar invariance across centre types.

Overall, changes in fit indices between consecutive models were minimal and fell clearly within the criteria established in the specialised literature. Consequently, the factorial structure of the MAILS is confirmed to be equivalent both by sex and by type of educational centre, although caution is warranted in the interpretation of invariance by institution type, given the marked group imbalance noted in Section 2.1.

3.3. Latent Profile Analysis

To address SO3 and identify differentiated AI literacy patterns, a Latent Profile Analysis (LPA) was conducted based on the mean scores of the nine MAILS dimensions. Progressive models from one to four profiles were estimated using maximum likelihood.

Optimal model selection was based on information criteria (AIC, BIC, SABIC), entropy, and class sizes, as shown in Table 10. Models were estimated under a class-invariant, diagonal covariance structure (tidyLPA Model 1, equivalent to the ‘EEI’ specification in mclust), in which item variances were constrained to be equal across classes and covariances among indicators were fixed to zero within each class, a parsimonious and frequently adopted specification for profile analysis with a moderate number of continuous indicators (Spurk et al., 2020). Models with five and six classes were additionally estimated and compared against the four-class solution using the Bootstrap Likelihood Ratio Test (BLRT) and posterior classification probabilities (see Section 3.3 for the rationale supporting the four-class solution).

Table 10. Comparison of latent profile analysis solutions.

No. Classes	LogLik	AIC	BIC	SABIC	Entropy	Min Class %	Max Class %	BLRT	BLRT
1	−20,043	40,123	40,210.79	40,154	—	100.0	100.0	—	—
2	−18,101	36,258	36,395.06	36,306	0.918	47.7	52.3	3885.0	<0.001
3	−17,263	34,603	34,788.03	34,667	0.934	14.0	47.5	1676.0	<0.001
4	−16,950	33,997	34,230.97	34,079	0.896	11.7	40.1	626.0	<0.001
5	−16,794	33,705	33,988	33,804	0.898	6.0	36.7	312.0	<0.001
6	−16,659	33,454	33,786	33,570	0.899	5.8	34.7	271.0	<0.001

Note. LogLik = log-likelihood; AIC = Akaike Information Criterion; BIC = Bayesian Information Criterion; SABIC = Sample-Size Adjusted BIC. Entropy reflects classification precision (values > 0.80 indicate good class differentiation). Model 1 (tidyLPA specification; equivalent to mclust ‘EEI’: diagonal covariance matrices, equal volume and shape across classes, covariances fixed to 0). BLRT = Bootstrap Likelihood Ratio Test comparing k vs. $k - 1$ classes (999 bootstrap draws); all BLRT tests were statistically significant ($p < 0.001$), including for class numbers beyond the selected solution. The four-class solution (shaded) was retained based on the elbow criterion (sharp reduction in marginal BIC/AIC improvement beyond $k = 4$; Δ BIC: $3 \rightarrow 4 = -557$, $4 \rightarrow 5 = -243$, $5 \rightarrow 6 = -202$), stable entropy (~ 0.90) across the 4–6 class range, and minimum class size, which fell below 6% of the sample ($n \approx 58$ and $n \approx 56$, respectively) for the 5- and 6-class solutions.

Comparison of models from one to four profiles showed progressive improvement in fit as the number of classes increased. To verify that the four-class solution was not arbitrarily selected, models with five and six classes were additionally estimated (Table 10). AIC, BIC, and SABIC continued to decrease through the six-class solution, and the Bootstrap Likelihood Ratio Test (BLRT, 999 bootstrap draws) was statistically significant at every step ($p < 0.001$), favouring each additional class over the previous solution, a pattern frequently observed with BLRT in large samples, where it tends to favour increasingly complex models and should therefore be interpreted jointly with parsimony, entropy, and class-size criteria rather than in isolation (Nylund et al., 2007). Following an elbow criterion, the marginal reduction in BIC dropped sharply beyond the four-class solution (Δ BIC: three-to-four classes = -557 ; four-to-five = -243 ; five-to-six = -202), indicating diminishing returns from additional classes. Entropy remained stable across the four-, five-, and six-class solutions (0.896, 0.898, and 0.899, respectively), showing no improvement in classification precision with additional classes; moreover, average posterior classification probabilities decreased for the five- and six-class solutions (minimum probability of 0.906 and 0.846, respectively, versus 0.914 for the four-class solution), indicating greater classification uncertainty. Finally, the five- and six-class solutions produced the smallest class comprising only 6.0% and 5.8% of the sample ($n \approx 58$ and $n \approx 56$, respectively), raising concerns about the substantive interpretability and stability of such small subgroups. Taken together, these criteria support retaining the four-class solution as the most parsimonious and substantively interpretable. Entropy was high across all solutions (> 0.80), indicating adequate precision in participant classification. Class proportions ranged from 11.7% to 40.1%, avoiding excessively small profiles, supporting the stability of the selected model.

The estimated means show a monotonically decreasing pattern across classes, evidencing a quantitative gradient of AI literacy. The “Very High” profile presents elevated scores across all dimensions, while the “Low” profile shows consistently reduced scores.

Intermediate profiles (High and Medium) reflect progressively lower levels. No qualitatively differentiated patterns between procedural and ethical dimensions were observed. Differentiation between profiles is fundamentally quantitative (see Table 11).

Table 11. Estimated dimension means by latent profile (four-class model).

Dimension	Very High (n = 166; 17.1%)	High (n = 302; 31.1%)	Medium (n = 389; 40.1%)	Low (n = 114; 11.7%)
Apply AI	8.54	7.21	5.07	2.45
Understand AI	8.93	7.29	5.15	2.47
Detect AI	8.88	7.33	4.98	2.31
AI Ethics	8.96	7.09	4.99	2.35
Create AI	7.23	5.25	4.20	2.40
Problem Solving	9.01	7.09	5.01	2.45
AI-Mediated Learning	8.65	6.20	4.49	2.27
Persuasion Literacy	8.85	7.38	5.65	2.65
Emotional Regulation	8.79	7.11	5.25	2.70

Note. Values correspond to estimated means from Latent Profile Analysis (LPA). Classes are ordered from highest to lowest overall AI literacy.

As shown in Figure 1, the identified profiles present a monotonically decreasing pattern across all AI literacy dimensions. The Very High profile shows consistently elevated scores, while the Low profile presents reduced levels across all dimensions. Intermediate profiles (High and Medium) are positioned progressively between both extremes, forming a continuous literacy gradient. The spatial overlap visible between these intermediate profiles in Figure 1 indicates that they are not fully separated in the latent space, which is consistent with the predominantly quantitative rather than qualitatively distinct nature of the identified gradient; membership in these intermediate profiles should therefore not be interpreted as reflecting sharply differentiated AI literacy types.

The spatial distribution of cases in the first two principal components reinforces this interpretation, with differentiation between profiles occurring primarily along the first component, which concentrates the largest proportion of explained variance. This pattern suggests that identified classes reflect fundamentally quantitative differences in overall AI literacy level rather than qualitatively distinct configurations across dimensions. Nevertheless, the degree of overlap observed between intermediate profiles indicates that differentiation is not entirely discrete, consistent with the continuous nature of the assessed construct.

Although the profiles primarily reflect a quantitative gradient, person-centred LPA classification offers two advantages over a simple total-score or percentile-based grouping. First, LPA jointly considers the relative configuration across all nine dimensions rather than collapsing them into a single composite, allowing detection of cases with comparatively uneven dimension profiles that a total score would mask; the residual overlap among intermediate classes reported above is itself diagnostic of this configurational variability, which a purely rank-based percentile split could not reveal. Second, LPA provides probabilistic classification with quantifiable uncertainty (entropy and posterior classification probabilities), whereas percentile-based cut-offs impose hard, arbitrary boundaries with no associated measure of classification confidence. These advantages are nonetheless modest given the predominantly gradient-like structure observed, and future studies should examine whether more qualitatively distinct AI literacy profiles emerge in samples with greater dimensional heterogeneity.

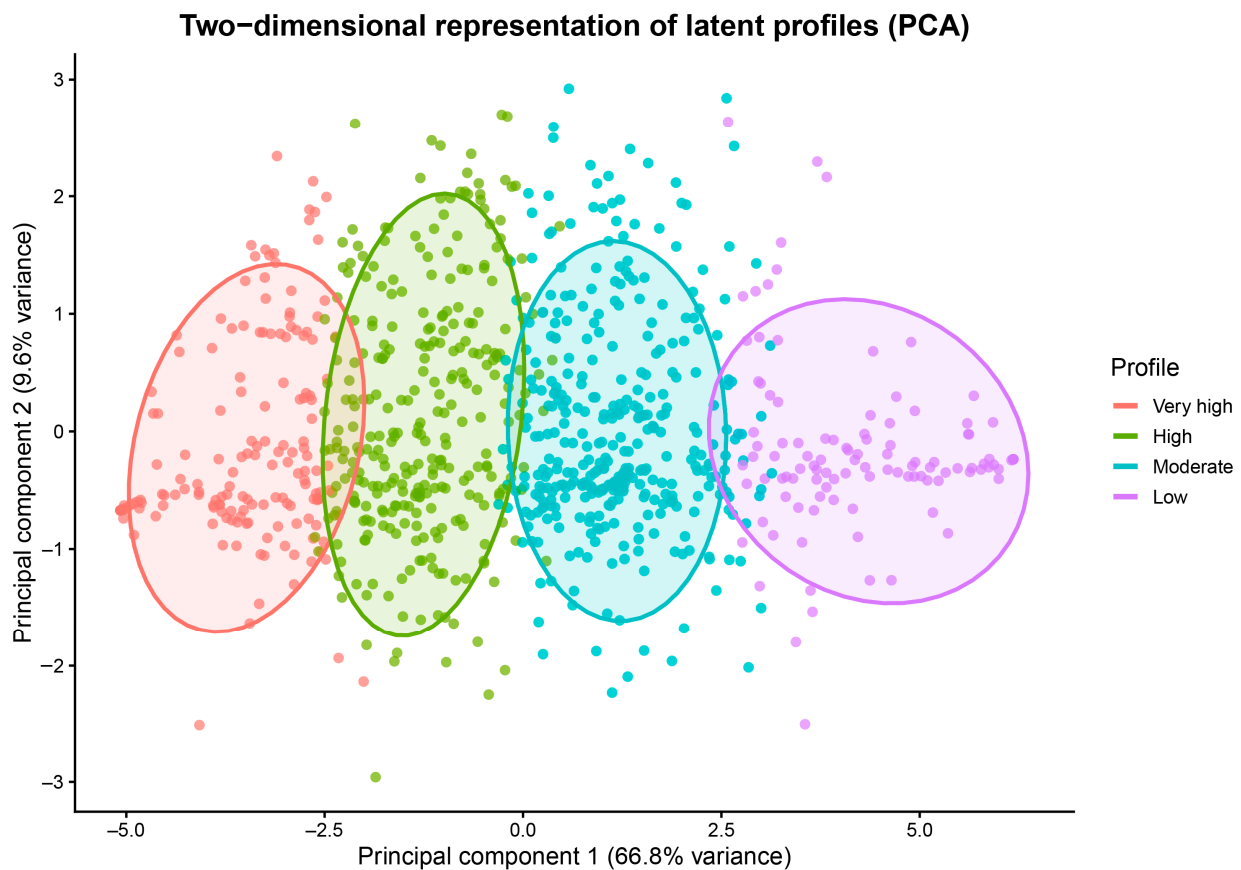


Figure 1. Two-dimensional representation of AI literacy latent profiles through principal component analysis.

3.4. Differences in Academic Internet Use by AI Literacy Profile

To examine whether differences in academic internet use existed across AI literacy profiles, a one-way ANOVA was conducted with profile as the independent variable and academic use as the dependent variable. Given the large sample size ($n = 971$), the procedure was considered robust to possible deviations from normality.

The analysis revealed statistically significant differences between profiles, $F(3, 967) = 8.824$, $p < 0.001$. The effect size was small ($\eta^2 = 0.027$), indicating that AI literacy profile accounts for approximately 2.7% of variance in academic internet use, a statistically significant but small effect. Means show an increasing trend in academic internet use as AI literacy level increases, with the Low profile presenting significantly lower values than all other profiles (see Table 12); this association should be interpreted as correlational given the cross-sectional design of the study.

Table 12. Differences in academic internet use by AI literacy profile.

Profile	N	M	SD
Low	114	2.009	1.101
High	302	2.427	1.081
Medium	389	2.473	1.004
Very High	166	2.669	1.178

Note. ANOVA: $F(3, 967) = 8.824$, $p < 0.001$, $\eta^2 = 0.027$, 95% CI [0.009, 0.048]. M = mean; SD = standard deviation.

A multinomial logistic regression was estimated to examine whether academic internet use and AI tool index predicted profile membership, using Very High as the reference category. Predictor variables were defined as follows. ACADEMIC_NET_USE is an ordi-

nal variable (range 1–5) representing daily internet use for academic purposes (1 = <2 h, 2 = 2–4 h, 3 = 4–6 h, 4 = 6–8 h, 5 = >8 h). AI_TOOL_INDEX is a composite index reflecting the diversity of AI tools used by each participant, calculated by summing the number of tool categories endorsed out of seven: translators (IA_TOOL_TRADUCTOR), chatbots (IA_TOOL_CHATBOT), writing assistants (IA_TOOL_ESCRITURA), analysis tools (IA_TOOL_ANALISIS), SEO/plagiarism detection tools (IA_TOOL_SEO_PLAGIO), video generation tools (IA_TOOL_VIDEO), and other AI tools (IA_TOOL_OTROS); higher scores indicate greater diversity of AI tool use (range 0–6, with each category coded 0 = not used, 1 = used). Profile membership was treated as the modal class assignment derived from the LPA posterior probabilities (hard assignment), without incorporating classification uncertainty into the regression model. This constitutes a limitation of the analysis, as the use of modal assignments may introduce downward bias in estimated associations; future studies should consider three-step approaches that account for classification uncertainty (Vermunt, 2010). Results are presented in Table 13.

Table 13. Multinomial logistic regression for the prediction of AI literacy profile. Reference category: Very High.

Comparison (vs. Very High)	Predictor	β	SE	z	p	OR	95% CI OR
High	ACADEMIC_NET_USE	−0.197	0.088	−2.240	0.025	0.82	[0.69, 0.97]
	AI_TOOL_INDEX	−0.056	0.108	−0.516	0.606	0.95	[0.77, 1.18]
Low	ACADEMIC_NET_USE	−0.599	0.125	−4.803	<0.001	0.55	[0.43, 0.70]
	AI_TOOL_INDEX	−0.331	0.151	−2.194	0.028	0.72	[0.53, 0.97]
Medium	ACADEMIC_NET_USE	−0.157	0.084	−1.874	0.061	0.85	[0.72, 1.00]
	AI_TOOL_INDEX	−0.063	0.103	−0.611	0.541	0.94	[0.77, 1.15]

Note. OR (Odds Ratio) = e^{β} . The 95% confidence intervals were calculated as $e^{(\beta \pm 1.96 \cdot SE)}$; CI = 95% confidence interval. $\chi^2(6) = 32.98$, $p < 0.001$ vs. null model.

Academic internet use showed a significant effect when comparing High and Low profiles against the Very High profile, substantially reducing the relative probability of belonging to the Low profile (OR = 0.55, $p < 0.001$). For the Medium profile, the effect was marginal ($p = 0.061$). The AI tool index only showed a significant effect when differentiating the Low profile from the Very High profile (OR = 0.72, $p = 0.028$), with no significant effects in other comparisons.

It should provide a concise and precise description of the experimental results, their interpretation, as well as the experimental conclusions that can be drawn.

4. Discussion

The present study aimed to validate the psychometric properties of the Meta AI Literacy Scale (MAILS) in a sample of 971 Ecuadorian high school students aged 14 to 19 years, as well as to explore the existence of differentiated AI literacy profiles and their relationship with the academic use of AI tools. Overall, the results obtained support the psychometric adequacy of the scale in this socio-educational context and provide empirical evidence that contributes to filling the gap in the literature on the validation of AI literacy instruments in Spanish-speaking school populations (Carolus et al., 2023; Ng et al., 2021; Zhong & Liu, 2025).

The central finding of this study is that AI literacy operates, in this sample, as an integrated and quantitatively continuous construct: the technical, ethical, and metacognitive dimensions do not show dissociated development profiles, but rather progress in an articulated manner along a competence gradient. This pattern is consistent with theoretical frameworks that conceive AI literacy as a multidimensional and indissociable construct, in which procedural mastery and critical awareness develop jointly (Chiu et al., 2024; Ng et al., 2021). The nine-factor correlated structure proposed by Carolus et al. (2023) showed

adequate fit in the analysed sample, with evidence of reliability, convergent and discriminant validity, and scalar invariance by sex and type of educational centre. The following sections discuss these findings in relation to the specialised literature and draw the most relevant implications for educational research and curricular policy design.

4.1. Factorial Structure, Reliability, and Validity of the MAILS

The nine-factor correlated structure originally proposed by Carolus et al. (2023) and confirmed in its abbreviated version by Koch et al. (2024) showed, in the present sample, fit indices within the ranges recommended in the structural equation modelling literature (Hu & Bentler, 1999; Kline, 2023). A direct comparison with previous validations is instructive. Carolus et al. (2023) validated the full MAILS with a German university student sample ($n = 1000$; $M = 24.6$ years; nine-factor solution; English-language items), obtaining CFI and factor loading values slightly below those reported here, which suggests that the factorial structure of the instrument generalises adequately to a younger (14–19 years), Spanish-speaking, secondary education population in Latin America. Koch et al. (2024) validated a 15-item abbreviated version of the MAILS with a similarly Western university sample, prioritising parsimony over dimensional breadth. Zhong and Liu (2025) validated a related AI literacy instrument with Chinese adolescents, finding a comparable multidimensional structure, though using a different number of dimensions and a different language version. Chiu et al. (2024) proposed an AI literacy assessment for Taiwanese secondary students. The present study differs from these antecedents in sample characteristics (secondary education, Ecuadorian context, marked institutional heterogeneity), language (Spanish Ecuadorian adaptation), and the explicit examination of measurement invariance by sex and institution type, an aspect absent from prior MAILS validations. These differences collectively strengthen the case for cross-cultural validation and justify the present contribution.

The evidence of convergent and discriminant validity reinforces this interpretation. AVE values above 0.50 across all dimensions indicate that items explain a substantial proportion of the variance of their respective latent factors, while HTMT indices below 0.90, and mostly below the conservative criterion of 0.85, support the empirical differentiation between conceptually close constructs, such as AI problem-solving self-efficacy and AI-mediated learning. It should be noted, however, that two HTMT values approached 0.89, near the upper bound of the liberal criterion, which indicates that discriminant validity between the corresponding factor pairs, while technically acceptable, is not robust; this proximity warrants caution against overstating the degree of factor independence. Following the framework of Ng et al. (2021), this pattern of results reaffirms that AI literacy constitutes a multidimensional and articulated construct, in which technical, ethical, and metacognitive components are distinguishable from each other without being independent, with direct implications for the design of assessment instruments and training programmes. These results also provide the empirical basis necessary to examine whether the identified structure is maintained equivalently across different sample groups, addressed in the following section.

4.2. Factorial Invariance of the MAILS by Sex and Type of Educational Centre

The results of the multigroup invariance analysis constitute one of the most methodologically relevant contributions of this study, insofar as previous research on the MAILS had not systematically examined measurement equivalence by sex or type of school funding (Carolus et al., 2023; Koch et al., 2024). The confirmation of configural, metric, and scalar invariance across both grouping variables—with CFI variations below 0.001, well below the recommended threshold of 0.010 (Cheung & Rensvold, 2002; Chen, 2007)—allows us to affirm that the scale items function equivalently regardless of participant sex and type of educational institution attended.

This finding has substantive implications at two levels. At the methodological level, scalar invariance supports the interpretation that differences in mean scores between groups are less likely to reflect measurement artefacts and more likely to reflect genuine variations in AI literacy level, legitimising comparisons between males and females and across different centre types. At the educational level, the result is especially pertinent in the Ecuadorian context, where the school system presents marked heterogeneity in terms of funding, technological resources, and students' socioeconomic characteristics. Measurement equivalence between state-funded and private institutions suggests that the MAILS captures the AI literacy construct coherently across school contexts with very different material and pedagogical conditions, expanding its diagnostic utility for comparative research and for the design of assessment policies in diversified educational systems. These results provide the empirical basis necessary for the latent profile analysis presented in the following section.

4.3. Latent Profiles of AI Literacy: Evidence of an Integrated Competence Gradient

Latent profile analysis identified four classes organised in a decreasing quantitative gradient, without any qualitatively atypical profile emerging that would evidence a dissociation between construct dimensions, such as would be expected if a student could show high procedural competence with limited ethical awareness, or vice versa. This result has direct theoretical relevance, as it challenges a modular view of AI literacy in which technical, ethical, and metacognitive dimensions could develop independently. On the contrary, the data are consistent with conceptual frameworks that conceive AI literacy as an integrated construct in which procedural mastery and critical reflection progress in an articulated manner (Chiu et al., 2024; Ng et al., 2021), and with the notion of progressive generalised competence that underlies the most consolidated digital literacy models in the international literature.

Differentiation between profiles occurs fundamentally along the first principal component, indicating that the four identified classes represent different levels of the same competence continuum rather than structurally divergent configurations. This gradient pattern has been observed in latent profile studies on digital competencies in adolescents (Zhong & Liu, 2025), suggesting that quantitative and hierarchical structuring may be a general property of digital literacy constructs, regardless of the instrument or cultural context employed. Overlap between intermediate profiles in the two-dimensional principal component analysis representation is, in this sense, an expected property consistent with the continuous nature of the construct and should not be interpreted as a limitation of the latent solution.

The four profiles can be characterised, in terms of their relative position in the competence gradient, as follows. The Very High profile (17.1%) groups students with elevated and balanced scores across all dimensions, including those of an ethical and metacognitive nature, indicating comprehensive mastery of AI literacy. The High profile (31.1%) encompasses users with high self-efficacy and solid systems understanding, although with somewhat more moderate values in the creation dimensions. The Medium profile (40.1%), which concentrates the largest sample proportion, presents functional literacy with more pronounced limitations in AI creation and AI-mediated learning dimensions. Finally, the Low profile (11.7%) shows reduced scores across the evaluated spectrum, indicating limited overall AI competence development; however, its precise characterisation as a group at risk of digital vulnerability would require complementary observational or longitudinal evidence that transcends the limitations of the present cross-sectional, self-reported design. The relationship between these profiles and academic AI use patterns is examined in the following section.

4.4. AI Literacy and Academic Use Patterns: Intentionality as an Associated Factor

Results from the ANOVA and multinomial logistic regression indicate that academic internet use is consistently associated with the level of AI literacy, with an increasing trend as the competence profile increases and a significant reduction in the probability of belonging to the Low profile when such use is more frequent ($OR = 0.55, p < 0.001$). This finding suggests that AI literacy does not operate solely as a declarative self-perception construct but rather articulates with differentiated academic practices. Given the cross-sectional study design, these associations must be interpreted in correlational rather than causal terms, without the possibility of establishing precedence relationships between academic use and competence level.

A complementary finding of interest is that the diversity of AI tools employed showed much more limited predictive capacity, with a significant effect only in the comparison between the Low and Very High profiles. This pattern suggests that the breadth of the technological repertoire does not constitute, by itself, an indicator of higher competence, and that the intentionality and academic purpose of use are more determinant than the mere variety of applications used. This result is consistent with the distinction, well documented in the digital literacy literature, between quantitative access to or use of technology and its qualitative cognitive appropriation (van Dijk, 2006; Zhong & Liu, 2025), and converges with warnings from international organisations regarding the risks of instrumental and unreflective use of AI in educational contexts (OECD, 2025b).

In this regard, the evidence provided by the OECD (2026) on the so-called “performance trap”, according to which AI-assisted students achieve superior results on practice tasks but show a decline in assessments conducted without technological assistance, offers a pertinent interpretive framework for contextualising this study’s findings; given that this evidence derives from a single experimental context rather than from consolidated cross-national research, the “performance trap” explanation should be treated as an interpretive hypothesis rather than as an established empirical fact, and its applicability to the Ecuadorian secondary school context remains to be empirically verified. Following Fuster Rabella (2025), this phenomenon emerges when AI is used as a substitute for the cognitive process rather than as an instrument for its amplification, which is especially relevant in a sample where 92.1% use conversational assistants, and access occurs predominantly through mobile phones. These conditions of use, predominantly immediate and functional, may favour patterns of AI interaction more oriented towards obtaining results than towards developing critical competencies, connecting with the concern for a second digital divide defined not by access to technology but by the quality and depth of its use (Hargittai, 2002; OECD, 2026). However, verification of this interpretive hypothesis exceeds the limits of the present design and constitutes a particularly relevant future research direction.

4.5. Educational Implications

The study findings suggest several tentative implications for curricular design and educational policy at the high school level, which should be regarded as data-informed hypotheses requiring further empirical verification rather than as direct conclusions from a cross-sectional self-report study. First, the identified quantitative gradient structure suggests that AI literacy should not be approached as a set of independent training modules, one technical, one ethical, one metacognitive, but rather as an integrated competence whose progressive development can be stimulated in an articulated manner. In line with the most recent international reference frameworks, this implies designing curricular pathways that prioritise critical analysis of AI-generated responses, understanding of their biases and limitations, and reflective regulation of generative tool use (OECD, 2025a; Park, 2025).

Second, the association between intentional academic internet use and higher competence profiles, correlational in nature and modest in magnitude, reinforces the need to pedagogically guide technological access, beyond guaranteeing device or connectivity availability. State-funded centres, which represent the majority of the sample, concentrate the greatest potential impact in equity terms, given that an improvement in the quality of technology use in these environments could reduce competence gaps between profiles. In this regard, it is especially relevant to consider that predominantly mobile phone access poses specific challenges for the development of dimensions such as AI system creation or complex problem-solving, requiring pedagogical mediation strategies adapted to students' real use conditions (OECD, 2026). Finally, the predominance of Year 3 high school students in the sample indicates that a significant proportion of students access higher education or the labour market with an already-consolidated AI literacy level, underscoring the urgency of formative interventions in earlier years and of specific teacher training that allows translating technological access into quality learning (OECD, 2025a; Rodríguez-Torres & Boillos-Pereira, 2026).

4.6. Limitations and Future Research

The present study has several limitations that must be considered when interpreting its results. First, the use of self-reported measures introduces the risk of social desirability bias: participants may overestimate their perceived competence or respond according to what they consider socially expected in the academic context. The MAILS evaluates self-perceived competence, which may differ from actual technical performance in controlled application situations; therefore, future studies should complement self-reports with observational or performance-based assessment measures to triangulate perception and execution (Markus et al., 2025). Second, the cross-sectional design prevents the establishment of causal relationships between the examined variables, so that the identified associations between academic use and competence profiles must be understood as correlational. Third, non-probabilistic convenience sampling limits the generalisation of results to the Ecuadorian educational system as a whole. Fourth, profile membership in the multinomial logistic regression was treated as the modal LPA class assignment (hard assignment), without incorporating classification uncertainty. This approach may introduce downward bias in estimated associations, and future studies should consider three-step methods that account for classification uncertainty (e.g., BCH approach; Vermunt, 2010).

These limitations point to several priority research directions. The most relevant is the development of longitudinal studies examining the stability of the identified latent profiles throughout the transition from high school to higher education, with the aim of detecting critical periods of competence stagnation or acceleration. It would also be of interest to compare AI literacy across students from different educational and cultural contexts, in order to evaluate the cross-cultural robustness of the observed competence gradient (Lintner, 2024; Zhong & Liu, 2025). Finally, the design of quasi-experimental studies evaluating the impact of specific GenAI training interventions would allow determining whether increased academic intentionality in AI use effectively translates into improved competence profiles, a hypothesis that this study's data suggest but cannot confirm.

5. Conclusions

This study provides psychometric evidence on the validity, reliability, and invariance of the Meta AI Literacy Scale (MAILS) in a sample of 971 Ecuadorian high school students, contributing to filling the gap in the literature on the validation of AI literacy instruments in Ecuadorian secondary education and offering a basis for future replications in other Spanish-speaking school contexts. Findings are articulated around three main contributions.

Instrument validity and measurement equity. Confirmatory factor analysis supported the nine-factor correlated structure originally proposed by Carolus et al. (2023), with fit indices within the ranges recommended in the specialised literature. Confirmation of scalar invariance by sex and type of educational centre, with CFI variations below 0.001, constitutes a result of particular methodological relevance, ensuring that comparisons between groups of different sex and institutional funding type are less likely to be affected by systematic measurement bias. This finding expands the applicability of the instrument in educational systems with marked socioeconomic and institutional heterogeneity, such as the Ecuadorian one.

Integrated and gradual nature of AI literacy. Latent profile analysis identified four classes, Very High, High, Medium, and Low, organised in a continuous quantitative gradient, without any profile emerging with dissociation between the technical and ethical dimensions of the construct. It is important to emphasise that these profiles are based entirely on students' self-perceived competence rather than on objective measures of AI literacy performance; the gradient should not be interpreted as reflecting actual AI literacy ability, but rather as reflecting how students perceive their own AI-related competencies. This result is consistent with theoretical frameworks that conceive AI literacy as an integrated multidimensional construct (Chiu et al., 2024; Ng et al., 2021) and suggests that, in this sample, procedural mastery and ethical awareness tend to develop in an articulated manner. Nevertheless, replication of this pattern in other cultural contexts and with performance measures, not only self-reported, is necessary before generalising this conclusion.

Association between use intentionality and competence level. Multinomial logistic regression analyses showed that academic internet use is significantly associated, though with a small effect size, with membership in higher AI literacy profiles, while the diversity of tools employed showed much more limited predictive capacity. This distinction between the quantity and purpose of technological use has direct implications for the design of formative interventions: access to multiple tools does not appear sufficient for the development of critical competencies if not accompanied by pedagogical mediation that guides their use towards reflective academic purposes. Given the cross-sectional study design, this association must be interpreted in correlational rather than causal terms.

Overall, the results support the utility of the MAILS as a valid and equitable instrument for evaluating AI literacy in the Ecuadorian secondary school context analysed; broader generalisation to other Spanish-speaking school systems requires further validation in different countries and cultural settings. The findings also identify the intentionality of academic technology use as a priority variable of interest for educational research and curricular design in secondary education. Validation of these findings through longitudinal designs and observational measures constitutes the most relevant line of continuity for future research.

Author Contributions: Conceptualization, D.C.T.-V., K.P.H.-M., Á.-F.R.-T. and J.-A.M.-M.; methodology, Á.-F.R.-T. and J.-A.M.-M.; software, J.-A.M.-M.; formal analysis, J.-A.M.-M.; investigation, K.P.H.-M., D.C.T.-V. and Á.-F.R.-T.; data curation, J.-A.M.-M.; writing—original draft preparation, K.P.H.-M. and D.C.T.-V.; writing—review and editing, Á.-F.R.-T. and J.-A.M.-M.; visualization, J.-A.M.-M.; supervision, Á.-F.R.-T. and J.-A.M.-M.; project administration, Á.-F.R.-T. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: This study was conducted in accordance with the Declaration of Helsinki and approved by the Ethics Committee of Central University of Ecuador (protocol code 001-PSN-2024, 27 February 2024).

Informed Consent Statement: Informed consent was obtained from the legal guardians of all minor participants involved in the study. All participants aged 18 or over provided their own informed consent.

Data Availability Statement: The data supporting the findings of this study are available from the corresponding author upon reasonable request.

Acknowledgments: During the preparation of this manuscript, the authors used ChatGPT (model 5.2, OpenAI) for the purposes of linguistic revision and editing of the manuscript text. The authors have reviewed and edited the output and take full responsibility for the content of this publication. The authors also wish to express their gratitude to the Research Group EducaTech-XXI (SEJ-666) “Education and Technology for the 21st Century” of the University of Granada, and to the Research Group Diversity, Health, Education and Sport (DISAED), code 045-GI-DI-2024, of the Universidad Central del Ecuador, for their support.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

MAILS	Meta AI Literacy Scale
AI	Artificial Intelligence
CFA	Confirmatory Factor Analysis
LPA	Latent Profile Analysis
CFI	Comparative Fit Index
TLI	Tucker–Lewis Index
RMSEA	Root Mean Square Error of Approximation
SRMR	Standardised Root Mean Square Residual
AVE	Average Variance Extracted
HTMT	Heterotrait–Monotrait ratio
AIC	Akaike Information Criterion
BIC	Bayesian Information Criterion
SABIC	Sample-Size Adjusted Bayesian Information Criterion
OECD	Organisation for Economic Co-operation and Development
ICT	Information and Communication Technologies
OR	Odds Ratio
SD	Standard Deviation
GenAI	Generative Artificial Intelligence
ANOVA	Analysis of Variance

Appendix A. Spanish (Ecuadorian) Adaptation of the Meta AI Literacy Scale (Carolus et al., 2023)

Table A1 presents the Spanish (Ecuadorian) adaptation of the MAILS, comprising 34 items distributed across nine dimensions, with an 11-point Likert response format (0 = “never or very rarely” to 10 = “very often”). The table includes the verbatim wording of each item, its corresponding dimension, and the variable code used in the database, in line with the adaptation procedure described in Section 2.2.

Table A1. Spanish (Ecuadorian) Adaptation of the Meta AI Literacy Scale.

Dimension	Code	No.	Item Statement
Aplicación y uso de la IA (Use & Apply AI)	MAILS_A1	1	Puedo manejar aplicaciones de IA en la vida cotidiana.
	MAILS_A2	2	Puedo utilizar aplicaciones de IA para facilitar mi vida cotidiana.
	MAILS_A3	3	Puedo utilizar la inteligencia artificial de forma significativa para alcanzar mis objetivos cotidianos.
	MAILS_A4	4	En la vida cotidiana, puedo interactuar con la IA de forma que facilite mis tareas.
	MAILS_A5	5	En la vida cotidiana, puedo colaborar provechosamente con una inteligencia artificial.
	MAILS_A6	6	Puedo comunicarme provechosamente con una inteligencia artificial en la vida cotidiana.
Comprensión de la IA (Understand AI)	MAILS_U1	7	Conozco los conceptos más importantes del tema «inteligencia artificial».
	MAILS_U2	8	Conozco definiciones de inteligencia artificial.
	MAILS_U3	9	Sé valorar cuáles son las limitaciones y oportunidades de utilizar una IA.
	MAILS_U4	10	Puedo evaluar qué ventajas e inconvenientes conlleva el uso de una inteligencia artificial.
	MAILS_U5	11	Puedo pensar en nuevos usos para la IA.
	MAILS_U6	12	Puedo imaginar posibles usos futuros de la IA.
Detección de la IA (Detect AI)	MAILS_D1	13	Puedo saber si estoy ante una aplicación basada en inteligencia artificial.
	MAILS_D2	14	Puedo distinguir los dispositivos que utilizan IA de los que no.
	MAILS_D3	15	Puedo distinguir si interactúo con una IA o con un «humano real».
Ética de la IA (AI Ethics)	MAILS_E1	16	Puedo sopesar las consecuencias del uso de la IA para la sociedad.
	MAILS_E2	17	Sé incorporar consideraciones éticas a la hora de decidir si utilizo los datos proporcionados por una IA.
	MAILS_E3	18	Soy capaz de analizar las implicaciones éticas de las aplicaciones basadas en la IA.
Creación de IA (Create AI)	MAILS_C1	19	Puedo diseñar nuevas aplicaciones de IA.
	MAILS_C2	20	Puedo programar nuevas aplicaciones en el campo de la «inteligencia artificial».
	MAILS_C3	21	Puedo desarrollar nuevas aplicaciones de IA.
	MAILS_C4	22	Puedo seleccionar herramientas útiles (por ejemplo, marcos de trabajo, lenguajes de programación) para programar una IA.

Table A1. *Cont.*

Dimension	Code	No.	Item Statement
Resolución de problemas con IA (AI Problem-Solving Self-Efficacy)	MAILS_PS1	23	Puedo confiar en mis habilidades en situaciones difíciles cuando utilizo IA.
	MAILS_PS2	24	Puedo manejar bien por mí mismo la mayoría de los problemas relacionados con la inteligencia artificial.
	MAILS_PS3	25	Por lo general, también puedo resolver bien las tareas extenuantes y complicadas cuando trabajo con inteligencia artificial.
Aprendizaje sobre IA (Learning)	MAILS_L1	26	Puedo estar al día de las últimas innovaciones en aplicaciones de IA.
	MAILS_L2	27	A pesar de los rápidos cambios en el campo de la inteligencia artificial, siempre puedo mantenerme al día.
	MAILS_L3	28	Aunque a menudo hay nuevas aplicaciones de IA, consigo estar siempre «al día».
Resistencia a la persuasión (Persuasion)	MAILS_P1	29	No dejo que la IA me influya en mis decisiones cotidianas.
	MAILS_P2	30	Puedo evitar que una IA me influya en mis decisiones cotidianas.
	MAILS_P3	31	Me doy cuenta de si la inteligencia artificial me influye en mis decisiones cotidianas.
Autorregulación emocional (Emotion Regulation)	MAILS_R1	32	Mantengo el control sobre sentimientos como la frustración y la ansiedad mientras hago cosas cotidianas con la IA.
	MAILS_R2	33	Puedo manejar cuando las interacciones cotidianas con la IA me frustran o me asustan.
	MAILS_R3	34	Puedo controlar la euforia que surge cuando utilizo la inteligencia artificial para fines cotidianos.

Appendix B. Descriptive Statistics, Skewness, and Kurtosis per Item

Table A2. Presents the mean, standard deviation, skewness, and kurtosis for each of the 34 MAILES items (n = 971), addressing the request for normality diagnostics to support the choice of estimator.

Factor (Dimension)	Item	M	SD	Skewness	Kurtosis
Factor 1 (Use and Application of AI)	MAILES_A1	6.181	2.811	−0.2608	−1.015
	MAILES_A2	6.081	2.775	−0.1798	−1.005
	MAILES_A3	5.890	2.696	−0.1165	−0.972
	MAILES_A4	6.342	2.764	−0.3229	−0.966
	MAILES_A5	6.036	2.675	−0.1355	−0.992
	MAILES_A6	5.699	2.738	−0.0102	−1.035
Factor 2 (Understanding of AI)	MAILES_U1	5.763	2.552	−0.0405	−0.831
	MAILES_U2	5.755	2.563	−0.0700	−0.917
	MAILES_U3	6.294	2.599	−0.2733	−0.830
	MAILES_U4	6.529	2.560	−0.3332	−0.785
	MAILES_U5	6.240	2.606	−0.2083	−0.867
	MAILES_U6	6.396	2.652	−0.2590	−0.924
Factor 3 (Detection of AI)	MAILES_D1	6.057	2.627	−0.1595	−0.886
	MAILES_D2	5.815	2.661	−0.0633	−0.927
	MAILES_D3	6.361	2.838	−0.2794	−1.057
Factor 4 (AI Ethics)	MAILES_E1	6.022	2.491	−0.1522	−0.699
	MAILES_E2	5.982	2.456	−0.1316	−0.701
	MAILES_E3	6.078	2.463	−0.1231	−0.775
Factor 5 (Creation of AI)	MAILES_C1	4.697	2.714	0.3510	−0.841
	MAILES_C2	4.784	2.715	0.2840	−0.920
	MAILES_C3	4.626	2.742	0.3280	−0.934
	MAILES_C4	5.272	2.721	0.1381	−0.935
Factor 6 (AI Self-Efficacy—Problem Solving)	MAILES_PS1	5.932	2.505	−0.0864	−0.807
	MAILES_PS2	5.904	2.477	−0.0711	−0.809
	MAILES_PS3	6.329	2.567	−0.2735	−0.827
Factor 7 (AI Self-Efficacy—AI-Mediated Learning)	MAILES_L1	5.449	2.622	0.0335	−0.859
	MAILES_L2	5.552	2.591	0.0367	−0.878
	MAILES_L3	5.450	2.683	0.0518	−0.960
Factor 8 (Own Competence—AI Persuasion Literacy)	MAILES_P1	6.371	2.973	−0.2981	−1.140
	MAILES_P2	6.556	3.034	−0.4101	−1.131
	MAILES_P3	6.257	2.926	−0.2796	−1.098
Factor 9 (Own Competence—Emotional Regulation in AI Interaction)	MAILES_R1	6.035	2.774	−0.1502	−1.007
	MAILES_R2	6.089	2.714	−0.1663	−0.975
	MAILES_R3	6.324	2.646	−0.2782	−0.861

Note. M = Mean; SD = Standard deviation.

Appendix C. Complete Standardised Factor Loadings, Standard Errors, Z Values, Significance, and R² for the Nine-Factor CFA Model

Table A3. Presents the complete standardised factor loadings, standard errors, Z values, statistical significance, and item-level R² for the nine-factor correlated CFA model (n = 971), addressing the request for a comprehensive psychometric reporting table.

Factor (Dimension)	Item	Standardised Loading (λ)	SE	Z	p	R ²
Factor 1 (Use and Application of AI)	MAILS_A1	0.8241	0.01650	49.94	<0.001	0.679
	MAILS_A2	0.8956	0.00863	103.80	<0.001	0.802
	MAILS_A3	0.8839	0.01064	83.06	<0.001	0.781
	MAILS_A4	0.8366	0.01537	54.42	<0.001	0.700
	MAILS_A5	0.8517	0.01489	57.21	<0.001	0.725
	MAILS_A6	0.8022	0.01842	43.54	<0.001	0.644
Factor 2 (Understanding of AI)	MAILS_U1	0.8105	0.01619	50.04	<0.001	0.657
	MAILS_U2	0.8374	0.01279	65.50	<0.001	0.701
	MAILS_U3	0.8556	0.01193	71.70	<0.001	0.732
	MAILS_U4	0.8430	0.01359	62.02	<0.001	0.711
	MAILS_U5	0.8282	0.01454	56.96	<0.001	0.686
	MAILS_U6	0.7881	0.01716	45.91	<0.001	0.621
Factor 3 (Detection of AI)	MAILS_D1	0.8366	0.01664	50.27	<0.001	0.700
	MAILS_D2	0.8515	0.01497	56.87	<0.001	0.725
	MAILS_D3	0.7854	0.01810	43.39	<0.001	0.617
Factor 4 (AI Ethics)	MAILS_E1	0.8476	0.01424	59.53	<0.001	0.718
	MAILS_E2	0.8758	0.01245	70.37	<0.001	0.767
	MAILS_E3	0.8828	0.01189	74.24	<0.001	0.779
Factor 5 (Creation of AI)	MAILS_C1	0.8880	0.01285	69.10	<0.001	0.789
	MAILS_C2	0.9240	0.00889	103.90	<0.001	0.854
	MAILS_C3	0.9276	0.00836	111.00	<0.001	0.860
	MAILS_C4	0.7538	0.01924	39.17	<0.001	0.568
Factor 6 (AI Self-Efficacy—Problem Solving)	MAILS_PS1	0.8721	0.01267	68.82	<0.001	0.761
	MAILS_PS2	0.8901	0.01164	76.48	<0.001	0.792
	MAILS_PS3	0.8472	0.01307	64.80	<0.001	0.718
Factor 7 (AI Self-Efficacy—AI-Mediated Learning)	MAILS_L1	0.8764	0.01350	64.91	<0.001	0.768
	MAILS_L2	0.8971	0.01331	67.42	<0.001	0.805
	MAILS_L3	0.8804	0.01355	64.99	<0.001	0.775
Factor 8 (Own Competence—AI Persuasion Literacy)	MAILS_P1	0.8366	0.01762	47.49	<0.001	0.700
	MAILS_P2	0.8796	0.01603	54.88	<0.001	0.774
	MAILS_P3	0.7895	0.02262	34.89	<0.001	0.623
Factor 9 (Own Competence—Emotional Regulation in AI Interaction)	MAILS_R1	0.8467	0.01733	48.86	<0.001	0.717
	MAILS_R2	0.8994	0.01234	72.85	<0.001	0.809
	MAILS_R3	0.8800	0.01387	63.46	<0.001	0.774

Note. Estimation by robust Maximum Likelihood (MLM), robust to non-normality. All loadings were statistically significant ($p < 0.001$). R² = item variance explained by its latent factor (standardised λ^2).

References

- Beaton, D. E., Bombardier, C., Guillemin, F., & Ferraz, M. B. (2000). Guidelines for the process of cross-cultural adaptation of self-report measures. *Spine*, 25(24), 3186–3191. [CrossRef] [PubMed]
- Bolaño-García, M., & Duarte-Acosta, N. (2024). Una revisión sistemática del uso de la inteligencia artificial en la educación [A systematic review of the use of artificial intelligence in education]. *Ética y Educación*, 39(1), 51–63. [CrossRef]
- Browne, M. W., & Cudeck, R. (1993). Alternative ways of assessing model fit. In K. A. Bollen, & J. S. Long (Eds.), *Testing structural equation models* (pp. 136–162). Sage.
- Buitrago, Á., García, A. M., & Torres-Ortiz, L. (2024). La alfabetización en inteligencia artificial: Propuesta articulada de dimensiones e indicadores [AI literacy: An articulated proposal of dimensions and indicators]. *Communication Papers—Media Literacy & Gender Studies*, 13(27), 118–139. Available online: <https://communicationpapers.revistes.udg.edu/article/view/23086/26652> (accessed on 12 March 2026).
- Carolus, A., Koch, M. J., Straka, S., Latoschik, M. E., & Wienrich, C. (2023). MAILS—Meta AI literacy scale: Development and testing of an AI literacy questionnaire based on well-founded competency models and psychological change- and meta-competencies. *Computers in Human Behavior: Artificial Humans*, 1(2), 100014. [CrossRef]
- Chen, F. F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural Equation Modeling: A Multidisciplinary Journal*, 14(3), 464–504. [CrossRef]
- Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling: A Multidisciplinary Journal*, 9(2), 233–255. [CrossRef] [PubMed]
- Chiu, T. K. F., Ahmad, Z., Ismailov, M., & Sanusi, I. T. (2024). What are artificial intelligence literacy and competency? A comprehensive framework to support them. *Computers and Education Open*, 6, 100171. [CrossRef]
- Clogg, C. C., & Bollen, K. A. (1991). Structural equations with latent variables. *Contemporary Sociology*, 20(1), 156. [CrossRef]
- Cruz-Pérez, M. A., Pozo-Vinueza, M. A., Aushay-Yupangui, H. R., & Arias-Parra, A. D. (2019). Las Tecnologías de la Información y de la Comunicación (TIC) como forma investigativa interdisciplinaria con un enfoque intercultural para el proceso de formación estudiantil [ICT as an interdisciplinary research approach with an intercultural perspective for student training]. *E-Ciencias de la Información*, 9(1), 1–15.
- Dúo Terrón, P., Moreno Guerrero, A. J., López Belmonte, J., & Marín Marín, J. A. (2023). Inteligencia Artificial y Machine Learning como recurso educativo desde la perspectiva de docentes en distintas etapas educativas no universitarias. *Revista interuniversitaria de investigación en tecnología educativa*, 58–78. [CrossRef]
- Fornell, C., & Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research*, 18(1), 39. [CrossRef]
- Fuster Rabella, M. (2025). *Evolving AI capabilities and the school curriculum: Emerging implications and a case study on writing* (OECD Education Working Papers No. 338). OECD Publishing. [CrossRef]
- Goth-Itoi, N. (2023, August 7). *Bringing AI literacy to high schools*. Stanford University. Available online: <https://hai.stanford.edu/news/bringing-ai-literacy-high-schools> (accessed on 12 March 2026).
- Guerrero-González, J., & Tendilla-Beltrán, E. R. (2022). Aprendizaje profundo: Redes neuronales [Deep learning: Neural networks]. *Nextia*, 15(8), 10–22.
- Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2019). *Multivariate data analysis* (8th ed.). Cengage.
- Hargittai, E. (2002). Second-level digital divide: Differences in people's online skills. *First Monday*, 7(4). [CrossRef]
- Henseler, J., Ringle, C. M., & Sarstedt, M. (2015). A new criterion for assessing discriminant validity in variance-based structural equation modeling. *Journal of the Academy of Marketing Science*, 43(1), 115–135. [CrossRef]
- Hernández-Nieto, R. A. (2002). *Contributions to statistical analysis: The coefficients of proportion of variance, content validity and kappa*. Universidad de Los Andes.
- Holmes, W., & Tuomi, I. (2022). State of the art and practice in AI in education. *European Journal of Education*, 57(4), 542–570. [CrossRef]
- Hu, L.-T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, 6(1), 1–55. [CrossRef]
- Hwang, G., Xie, H., Wah, B. W., & Gašević, D. (2020). Vision, challenges, roles and research issues of artificial intelligence in education. *Computers and Education: Artificial Intelligence*, 1, 100001. [CrossRef]
- International Test Commission. (2017). *The ITC guidelines for translating and adapting tests* (2nd ed.). Available online: https://www.intestcom.org/files/guideline_test_adaptation_2ed.pdf (accessed on 20 March 2026).
- JASP Team. (2025). *JASP* (version 0.96) [Computer software]. Available online: <https://jasp-stats.org/> (accessed on 20 April 2026).
- Kline, R. B. (2023). *Principles and practice of structural equation modeling* (5th ed.). Guilford Press.
- Koch, M. J., Carolus, A., Wienrich, C., & Latoschik, M. E. (2024). Meta AI literacy scale: Further validation and development of a short version. *Heliyon*, 10(22), e39686. [CrossRef] [PubMed]
- Lampropoulos, G., López-Belmonte, J., Pozo-Sánchez, S., & Marín-Marín, J.-A. (2026). Agentic AI in education. In S. Papadakis (Ed.), *Teacher perspectives and responsible practice for integrating AI in the classroom* (pp. 1–28). IGI Global Scientific Publishing. [CrossRef]

- Licona-Quiterio, O. (2024). IA y educación en México: Entre oportunidades y desafíos [AI and education in Mexico: Between opportunities and challenges]. *Revista Internacional de Estudios sobre Sistemas Educativos*, 3(15), 675–688.
- Lintner, T. (2024). A systematic review of AI literacy scales. *Science of Learning*, 9(1), 50. [CrossRef] [PubMed]
- López de Mántaras, R., & Brunet-Crosa, P. (2023). ¿Qué es la inteligencia artificial? [What is artificial intelligence?]. *Papeles de Relaciones Ecosociales y Cambio Global*, 164, 13–21.
- MacCallum, R. C., Browne, M. W., & Sugawara, H. M. (1996). Power analysis and determination of sample size for covariance structure modeling. *Psychological Methods*, 1(2), 130–149. [CrossRef]
- Markus, A., Carolus, A., & Wienrich, C. (2025). Objective measurement of AI literacy: Development and validation of the AI Competency Objective Scale (AICOS). *arXiv*, arXiv:2503.12921.
- Mora, M., Ochoa, C., Cango, M., & Gutiérrez, J. (2024). Innovación educativa en la universidad: Uso de TIC e inteligencia artificial para mejorar la enseñanza y evaluación [Educational innovation at university: Use of ICT and AI to improve teaching and assessment]. *Reincisol*, 3(6), 6409–6427. [CrossRef]
- Ng, D. T. K., Leung, J. K. L., Chu, S. K. W., & Qiao, M. S. (2021). Conceptualizing AI literacy: An exploratory review. *Computers and Education: Artificial Intelligence*, 2, 100041. [CrossRef]
- Nylund, K. L., Asparouhov, T., & Muthén, B. O. (2007). Deciding on the number of classes in latent class analysis and growth mixture modeling: A monte carlo simulation study. *Structural Equation Modeling*, 14(4), 535–569. [CrossRef]
- OECD. (2025a). *Results from TALIS 2024: The state of teaching*. OECD Publishing. [CrossRef]
- OECD. (2025b). *Empowering learners for the age of AI: An AI literacy framework for primary and secondary education (review draft)*. OECD Publishing. Available online: <https://ailiteracyframework.org> (accessed on 9 April 2026).
- OECD. (2026). *OECD digital education outlook 2026: Exploring effective uses of generative AI in education*. OECD Publishing. [CrossRef]
- Paredes-Gallardo, C. (2024). Aplicación de la inteligencia artificial en el ámbito educativo: Análisis de buenas prácticas y recomendaciones [Application of AI in education: Analysis of good practices and recommendations]. *Revista de Educación y Derecho*, 2, 512–539. [CrossRef]
- Park, J. (2025). A systematic literature review of generative artificial intelligence (GenAI) literacy in schools. *Computers and Education: Artificial Intelligence*, 9, 100487. [CrossRef]
- Pozo-Sánchez, S., Marín-Marín, J.-A., Rodríguez-Torres, Á.-F., & López-Belmonte, J. (2026). Inteligencia artificial en educación: Un estado de la cuestión en el contexto hispanoamericano. *Texto Livre* 19, e63138. [CrossRef]
- R Core Team. (2024). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. Available online: <https://www.R-project.org/> (accessed on 18 March 2026).
- Rodríguez-Torres, A. F., & Boillos-Pereira, M. M. (2026). Artificial intelligence from the point of view of university students. *European Public & Social Innovation Review*, 11, 1–21. [CrossRef]
- Rosenberg, J., Beymer, P., Anderson, D., van Lissa, C. J., & Schmidt, J. (2018). tidyLPA: An R package to easily carry out latent profile analysis (LPA) using open-source or commercial software. *Journal of Open Source Software*, 3(30), 978. [CrossRef]
- Schwartz, H. L., & Diliberti, M. K. (2026). *More students use AI for homework, and more believe it harms critical thinking: Selected findings from the American youth panel*. RAND Corporation. [CrossRef]
- Scrucca, L., Fop, M., Murphy, T. B., & Raftery, A. E. (2016). mclust 5: Clustering, classification and density estimation using Gaussian finite mixture models. *The R Journal*, 8(1), 289–317. [CrossRef]
- Shipley, B., & Douma, J. C. (2021). Generalized AIC and chi-squared statistics for path models consistent with directed acyclic graphs. *Ecology*, 102(3), e03304. [CrossRef] [PubMed]
- Spurk, D., Hirschi, A., Wang, M., Valero, D., & Kauffeld, S. (2020). Latent profile analysis: A review and “how to” guide of its application within vocational behavior research. *Journal of Vocational Behavior*, 120, 103445. [CrossRef]
- Suárez-Estavillo, U. (2025). La inteligencia artificial en la educación: ¿transformación o infoxicación? [AI in education: Transformation or information overload?]. *Sintaxis*, 14, 69–88. [CrossRef]
- Suárez-Muñoz, F., & Bautista-Tejeda, A. A. (2024). Autonomía de la inteligencia artificial: Una reflexión crítica desde la filosofía kantiana [Autonomy of artificial intelligence: A critical reflection from Kantian philosophy]. *SUMMA*, 6(1), 1–11. [CrossRef]
- Tamara-Peña, J. J., & García-López, R. I. (2024). Alfabetización inicial con inteligencia artificial [Initial literacy with artificial intelligence]. *Pedagogical Constellations*, 3(2), 115–122.
- UNESCO. (n.d.-a). *Artificial intelligence*. Available online: <https://www.unesco.org/en/artificial-intelligence> (accessed on 14 March 2026).
- UNESCO. (n.d.-b). *Artificial intelligence in education*. Available online: <https://www.unesco.org/en/digital-education/artificial-intelligence> (accessed on 12 March 2026).
- UNESCO. (2024). *AI in education: Making it work for people*. Available online: <https://www.unesco.org/en/digital-education> (accessed on 15 April 2026).
- van Dijk, J. A. G. M. (2006). Digital divide research, achievements and shortcomings. *Poetics*, 34(4–5), 221–235. [CrossRef]
- Vermunt, J. K. (2010). Latent class modeling with covariates: Two improved three-step approaches. *Political Analysis*, 18(4), 450–469. [CrossRef]

- Vieriu, A. M., & Petrea, G. (2025). The impact of artificial intelligence (AI) on students' academic development. *Education Sciences*, 15(3), 343. [CrossRef]
- Vodafone Foundation. (2025). *Young people and generative AI: Survey of 7000+ students aged 12–17 across seven European countries*. Vodafone Foundation. Available online: <https://vodafonefoundation.org> (accessed on 30 April 2026).
- Wickham, H. (2016). *ggplot2: Elegant graphics for data analysis* (2nd ed.). Springer International Publishing.
- Zhong, B., & Liu, X. (2025). Evaluating AI literacy of secondary students: Framework and scale development. *Computers & Education*, 227, 105230. [CrossRef]

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