International Journal of Human–Computer Interaction

ISSN: 1044-7318 (Print) 1532-7590 (Online) Journal homepage: www.tandfonline.com/journals/hihc20

Psychometric Evaluation of the Turkish Version of the Artificial Intelligence Self-Efficacy Scale Among University Students

Seda Güney & Remziye Semerci Şahin

INTERNATIONAL JOURNAL OF HUMAN–COMPUTER INTERACTION
https://doi.org/10.1080/10447318.2026.2626817

Taylor & Francis
Taylor & Francis Group

Check for updates

Psychometric Evaluation of the Turkish Version of the Artificial Intelligence Self-Efficacy Scale Among University Students

Seda Güney and Remziye Semerci Şahin

School of Nursing, Koç University, Istanbul, Turkey

ABSTRACT

University students constitute a key population for assessing competence and confidence in the use of artificial intelligence (AI); however, culturally adapted and psychometrically sound instruments to measure AI self-efficacy in Türkiye are limited. This methodological study aimed to translate, culturally adapt, and evaluate the psychometric properties of the Artificial Intelligence Self-Efficacy Scale (AISES) for Turkish university students. Data were collected between May and November 2025 via an online survey administered to 284 students from multiple universities using convenience sampling. Content validity was assessed through expert evaluation. Construct validity was examined using Exploratory and Confirmatory Factor Analyses conducted on randomly split subsamples. The analyses supported a four-factor structure explaining 73.69% of the variance, with acceptable model fit indices and excellent internal consistency (Cronbach's $\alpha = 0.937$). The Turkish AISES demonstrated strong validity and reliability, supporting its use in assessing AI self-efficacy and informing AI-related educational initiatives among university students in Türkiye.

KEYWORDS

Artificial intelligence; self-efficacy; validity; reliability; scale adaptation

1. Introduction

Artificial intelligence (AI) has rapidly become an integral part of education, healthcare, and professional practice, requiring individuals to develop not only technical skills but also confidence in using AI effectively. As AI technologies increasingly influence decision-making, learning environments, and service delivery, understanding and measuring individuals' perceived capability has become a crucial area of inquiry (Falebita & Kok, 2025; Luo & Hsiao-Chin, 2023). AI has become a defining force in higher education, transforming the way students learn, communicate, and solve problems (Chu et al., 2022; Stracke et al., 2025). AI technologies, such as adaptive learning platforms, intelligent tutoring systems, and automated feedback tools, are increasingly being incorporated into university curricula (Owan et al., 2023). As these technologies evolve, students are expected to not only use AI tools but also to understand, evaluate, and confidently interact with them (Bećirović et al., 2025). This growing demand underscores the importance of AI self-efficacy, which refers to students' confidence in their ability to effectively use and adapt to AI technologies (Falebita & Kok, 2025).

Self-efficacy refers to one's confidence in their capacity to perform specific tasks (Chen et al., 2001). Recent research highlights that learners' self-efficacy is closely interacted to emotional, motivational, and engagement-related processes in technology-enhanced learning environments. Wang et al. (2024) demonstrated that achievement emotions significantly shape learners' engagement and willingness to communicate, indicating that confidence in learning is embedded within affective experiences. Similarly, Yang et al. (2025) reported strong reciprocal relationships among motivation, engagement, and higher-order cognitive outcomes, emphasizing the central role of motivational beliefs in sustaining learning behaviors. These findings suggest that self-efficacy in AI-enhanced education should be conceptualized as a multidimensional construct influenced by emotional, motivational, and behavioral factors. Within educational contexts, AI self-efficacy has been identified as an important determinant of

CONTACT Remziye Semerci Şahin rsemerci@ku.edu.tr Koç Üniversitesi Health Sciences Campus, Davutpasa Cd. No:4 Topkapi, Istanbul, Turkey

technology acceptance, engagement, and readiness for innovation (Wang & Chuang, 2024). To assess this emerging construction, several instruments have been developed and validated across disciplines. The Artificial Intelligence Self-Efficacy Scale (AISES), originally developed by Wang and Chuang (2024), provided a psychometrically robust tool with four subscales. Besides this measure, there have also been other tools developed to measure the relationship between humans and AI, including the General Self-Efficacy Scale for Artificial Intelligence (GSE-6AI) (Morales-García et al., 2024), the Teacher AI Competence Self-Efficacy (TAICS) Scale (Chiu et al., 2025), and attitude-based tools such as the AI Attitude Scale (AIAS-4) (Grassini, 2023). These studies collectively emphasize the centrality of self-efficacy in shaping perceptions and readiness for AI-enhanced education.

In Türkiye, parallel research efforts have expanded rapidly. Several instruments have been translated and validated to assess AI-related perceptions, including the Artificial Intelligence Literacy Scale (Uluğ et al., 2025), the Generative AI Acceptance Scale (Yılmaz et al., 2023), the AI Anxiety Scale (Terzi, 2020), and most recently, the Turkish Adaptation of the AI Self-Efficacy Scale (Aca et al., 2025; Uyan & Gültekin, 2024). These studies have substantially advanced the measurement of AI-related constructs in Turkish contexts, confirming that the original scales generally retain valid and reliable structures. However, existing Turkish adaptations of AI self-efficacy measures have generally been conducted in limited or specialized populations, such as employees or mixed adult samples, and often lack strong educational context specificity. Given that psychometric properties are inherently context-dependent, validation within higher education populations, particularly among university students who are both active learners and future professionals expected to integrate AI into their professional practice, is essential to ensure construct stability and cross-context robustness. Studies have demonstrated that AI self-efficacy is closely associated with emotional, cognitive, and behavioral variables in educational settings, including AI anxiety, AI self-competence, student attitudes, and AI literacy. For instance, Asio and Sardina (2025) showed that AI self-efficacy indirectly influences AI anxiety through AI self-competency, highlighting the role of self-beliefs in shaping emotional responses to AI. Similarly, Albino et al. (2025) reported that AI anxiety and learning-related configurations significantly predict students' AI self-efficacy, emphasizing the importance of contextual learning experiences. In addition, Asio (2025) demonstrated that AI self-efficacy mediates the relationship between students' attitudes toward AI and AI literacy, underscoring its central role in AI-related learning processes. In Türkiye, previous studies have also emphasized the dynamic relationships between AI self-efficacy, digital literacy, and academic readiness (Atik & Erdemir, 2025; Eminoğlu & Çelikkanat, 2024); however, evidence remains limited regarding the psychometric performance of AI self-efficacy instruments specifically within university student populations and formal educational contexts. Therefore, the present study addresses this gap by providing a rigorous psychometric evaluation of the Turkish version of the Artificial Intelligence Self-Efficacy Scale among university students. This contribution to the literature offers a context-sensitive, methodologically robust measurement tool for future educational research and intervention studies aimed at strengthening AI-related competencies in higher education.

Therefore, the present study aims to assess the validity and reliability of the Turkish version of the Artificial Intelligence Self-Efficacy Scale among university students. By focusing on this population, the study aims to verify the factor structure, internal consistency, and construct validity of the instrument within the context of higher education. Establishing a robust, education-specific Turkish measure of AI self-efficacy will support future research, curriculum development, and policy initiatives aimed at fostering AI literacy and competence in universities.

2. Methods

2.1. Study design

This research employed a descriptive methodological design aimed at adapting and validating the Artificial Intelligence Self-Efficacy Scale for use among Turkish university students. The study followed a systematic, multistage process encompassing translation, cultural adaptation, and psychometric evaluation, consistent with the STROBE (Strengthening the Reporting of Observational Studies in

Epidemiology) guidelines to ensure methodological transparency, completeness, and replicability in reporting (Von Elm et al., 2007).

2.2. Study setting

This study was conducted among university students in Türkiye between May and November 2025. Participants were recruited from multiple universities, including both public and private institutions, across different geographical regions of the country. Data were collected using an online, structured questionnaire distributed via social media platforms, including Instagram, X, and WhatsApp, as well as academic groups. This approach facilitated access to students from diverse academic programs and institutional contexts, thereby increasing heterogeneity within the study population and supporting broader representation of university students in Türkiye.

2.3. Study sample

The sample size for this methodological study was determined in accordance with established recommendations for scale adaptation and psychometric validation research. As emphasized by Gunawan et al. (2021), there is no single, universally accepted rule for determining sample size in factor analytic studies. However, larger samples generally enhance factor stability, reduce measurement error, and improve the generalizability of findings. The authors summarized methodological conventions in the literature, suggesting that item-to-participant ratios typically range from 1:5 to 1:10 for Exploratory Factor Analysis (EFA) and Confirmatory Factor Analysis (CFA). This guidance is consistent with the recommendations of Hair et al. (2010), who argue that a minimum of five to ten participants per item is appropriate for psychometric analyses. Based on these established standards, the current study adopted the upper boundary of this recommendation, applying a 1:10 ratio to the 22-item Artificial Intelligence Self-Efficacy Scale (AISES). Consequently, a minimum of 220 participants was considered methodologically adequate to ensure reliable factor extraction and stable parameter estimation. Ultimately, the study was completed with 284 university students, exceeding the recommended minimum sample size and thereby strengthening the statistical power and robustness of the validity and reliability analyses conducted on the Turkish version of the scale (Gunawan et al., 2021).

The inclusion criteria for this study were being enrolled as a student at a university, having proficiency in Turkish, and voluntarily agreeing to participate in the research. Students who met these criteria were invited to complete the online survey. The exclusion criteria included submitting incomplete data collection forms or withdrawing from the study at any stage. Participants who failed to complete the questionnaire in its entirety were excluded from the final analysis to ensure data quality and consistency.

2.4. Instruments

2.4.1. Sociodemographic Information Form

The Form was developed by researchers to collect data on participants' personal and educational characteristics. The form included questions regarding age, gender, field of study, year of study, previous experience with artificial intelligence technologies, and frequency of technology use.

2.4.2. Artificial Intelligence Self-Efficacy Scale (AISES)

The AISES was developed to evaluate individuals' self-efficacy perceptions regarding the use of artificial intelligence technologies (Wang & Chuang, 2024). The scale consists of a total of 22 items and includes four subdimensions: Assistance, Anthropomorphic Interaction, Comfort with AI, and Technological Skills. Items are rated on a 7-point Likert scale, with response options ranging from "strongly disagree" to "strongly agree." Exploratory and confirmatory factor analyses conducted as part of the validity and reliability assessment demonstrated that the four-factor structure exhibited a good model fit (CFI = 0.941, RMSEA = 0.079). The overall internal consistency of the scale was high, with a Cronbach's alpha coefficient of 0.958. The alpha coefficients for the subdimensions ranged between 0.869 and 0.970.

2.5. Data collection process

Data collection was carried out between May 2025 and November 2025 using a structured online questionnaire developed in Google Forms. The research team disseminated the survey link through widely used social media platforms, including Instagram, WhatsApp, and X (Twitter), to reach university students from various regions of Türkiye. This digital approach enabled efficient and geographically diverse recruitment of participants enrolled in different universities. The first page of the survey presented an information and informed consent form, which detailed the study's objectives, principles of voluntary participation, and assurances of data confidentiality. Participants who provided consent by selecting the agreement checkbox were granted access to the main sections of the questionnaire. The instrument consisted of two components: the Sociodemographic Information Form, designed to capture demographic and educational characteristics relevant to the research objectives, and the AISES, which assessed participants who perceived self-efficacy in using artificial intelligence technologies. Participants were informed that they could discontinue participation at any point without any justification or consequence. Data was collected anonymously, and all responses were securely stored in encrypted digital files with restricted access limited to the research team. The average time required to complete the survey was approximately 10 min.

2.6. Data Analysis process

All statistical analyses were carried out using IBM SPSS Statistics version 28.0 and AMOS version 24.0 software. Descriptive statistics, including means, standard deviations, frequencies, and percentages, were computed to summarize participants' sociodemographic characteristics and responses to AI-related variables. Content validity was examined through expert evaluation, and the Item-Level Content Validity Index (I-CVI) and Scale-Level Content Validity Index (S-CVI) were calculated. Construct validity was assessed using a two-step approach that combined Exploratory Factor Analysis (EFA) and Confirmatory Factor Analysis (CFA). The dataset was randomly divided into two subsamples: one for EFA and the other for CFA. EFA was conducted using principal component analysis with varimax rotation to identify the underlying factor structure of the Turkish version of the Artificial Intelligence Self-Efficacy Scale (AISES). The suitability of the data for factor analysis was confirmed through the Kaiser Meyer Olkin (KMO) measure and Bartlett's Test of Sphericity, which indicated excellent sampling adequacy and factorability. Subsequently, CFA was performed to validate the four-factor structure identified by the EFA. Model fit was evaluated using multiple indices, including chi-square (χ^2), degrees of freedom (df), χ^2/df ratio, Root Mean Square Error of Approximation (RMSEA), Goodness of Fit Index (GFI), Comparative Fit Index (CFI), Relative Fit Index (RFI), Incremental Fit Index (IFI), Normed Fit Index (NFI), and Tucker–Lewis Index (TLI). The internal consistency reliability of the scale and its subdimensions was assessed using Cronbach's alpha coefficients and item–total correlations. In addition, split-half reliability was calculated using the Spearman–Brown and Guttman split-half coefficients to evaluate the scale's stability over item halves. Finally, Hotelling's T^2 test was applied to detect potential response bias and to examine whether participants provided consistent responses across items. The test yielded statistically significant results, confirming the absence of uniform response patterns and supporting the scale's discriminative capacity. All analyses were conducted at a 95% confidence interval, with the significance level set at $p < 0.05$.

2.6.1. Language validity

The linguistic equivalence of the scale was ensured through a rigorous translation and back-translation procedure. Initially, three independent bilingual experts whose native language was Turkish translated the original English version into Turkish (Çapık et al., 2018). The research team then reviewed all versions, comparing semantic and conceptual consistency, and produced a unified draft by consensus. Subsequently, two bilingual linguists, fluent in both Turkish and English, back-translated this version into English. The back-translated form was carefully compared with the original instrument to identify any conceptual discrepancies or loss of meaning (Yasir, 2016). Necessary adjustments were made until full linguistic equivalence was achieved. This translation process followed the World Health Organization (2017) guidelines for the cross-cultural adaptation of research instruments.

2.6.2. Content validity

The Turkish version of the AISES underwent a comprehensive expert assessment to evaluate its content validity following the completion of linguistic adaptation. Eight domain experts with advanced qualifications in nursing, artificial intelligence, and educational technologies independently reviewed each item, focusing on its clarity, conceptual relevance, and representational adequacy within the cultural context. Experts rated each statement on a four-point scale (1 = not relevant, 4 = highly relevant) as recommended by Polit et al. (2007). Based on these ratings, both the Item-Level Content Validity Index (I-CVI) and the Scale-Level Content Validity Index (S-CVI) were calculated. Items with I-CVI values below 0.80 were revised for clarity or conceptual appropriateness.

2.6.3. Preliminary study

A preliminary evaluation was conducted to ensure the linguistic clarity, conceptual relevance, and cultural appropriateness of the Turkish version of the scale. This phase followed the pilot application stage recommended in intercultural adaptation frameworks, emphasizing the assessment of comprehensibility and cultural equivalence among the target population. Preliminary testing was conducted with 30 university students who met the inclusion criteria but were not part of the main sample. Participants were asked to complete the scale and provide feedback regarding the clarity, comprehensibility, and cultural suitability of each item. In line with the recommendations of the World Health Organization (2017) and Çapık et al. (2018), participants were also encouraged to comment on potentially ambiguous, culturally inappropriate, or difficult-to-understand terms. Feedback from this stage informed minor linguistic and contextual revisions, which aimed to enhance semantic and conceptual equivalence between the original and Turkish versions. Data obtained from this stage were excluded from the main statistical analyses, as the purpose was to confirm the comprehensibility and cultural relevance of the adapted instrument before conducting its psychometric evaluation.

2.6.4. Construct validity

Construct validity was examined through a combination of Exploratory Factor Analysis (EFA) and Confirmatory Factor Analysis (CFA) using two independent subsamples. Sampling adequacy was verified through the Kaiser–Meyer–Olkin (KMO) measure and Bartlett's Test of Sphericity, which confirmed the suitability of the data for factor analysis (KMO > 0.80; $p < 0.001$). EFA was conducted using Principal Axis Factoring (PAF) with varimax rotation, following the recommendations of Howard (2016) and Costello and Osborne (2005), who emphasized the advantages of PAF over principal component analysis for identifying latent constructs. Factor retention was guided by eigenvalues greater than 1, the scree plot, and parallel analysis, in accordance with contemporary best practices. Factor loadings were interpreted using the 0.40–0.30–0.20 rule suggested by Howard (2016), where primary loadings ≥ 0.40, cross-loadings ≤ 0.30, and loading differences ≥ 0.20 were considered satisfactory indicators of construct distinctiveness. The factor structure derived from EFA accounted for an acceptable proportion of the total variance (>60%), supporting the theoretical dimensionality of the scale.CFA was subsequently performed on the second subsample using AMOS version 24.0 to verify the factor structure identified in the exploration phase. Model adequacy was assessed using multiple goodness-of-fit indices, including χ^2/df, RMSEA, GFI, CFI, TLI, and IFI. Model fit thresholds were interpreted according to Hair et al. (2020), where $\chi^2/df < 5$, RMSEA < 0.08, and CFI, TLI, and IFI ≥ 0.90, along with GFI ≥ 0.90, indicated an acceptable model fit. The confirmatory results demonstrated that the proposed factor structure achieved satisfactory fit indices, confirming the construct validity and structural coherence of the Turkish version of the scale.

2.6.5. Reliability analysis

The internal consistency of the scale was assessed through Cronbach's alpha coefficients for the total scale and its subdimensions. Additionally, item–total correlation coefficients, split-half reliability (Spearman–Brown and Guttman coefficients) (Eisinga et al., 2013), and Hotelling's T^2 tests were calculated to confirm the reliability and internal consistency of the instrument (Şencan, 2005). Cronbach's alpha values exceeding 0.70 indicated satisfactory internal consistency across all subscales (Hajjar, 2018). The results collectively supported the psychometric robustness and measurement stability of the Turkish version of the AISES.

2.7. Ethical considerations

This study was conducted in accordance with the ethical principles outlined in the Declaration of Helsinki and relevant national research ethics guidelines. Ethical approval was obtained from the Koç University Ethics Committee (Approval No: 2025.224.IRB3.085) prior to data collection. Institutional permission was also secured from the university administration. Participation in the study was voluntary, and no incentives were provided. Before data collection, all participants received detailed information regarding the purpose, procedures, potential benefits, and confidentiality of the study. Informed consent was obtained electronically from each participant, who was also reminded of their right to withdraw from the study at any time without penalty. All data was collected anonymously through an online survey platform, and no personally identifiable information was stored.

3. Results

3.1. Sample characteristics

Most participants were women (76.1%) and were enrolled in the nursing department (92.6%). The mean age of the students was 21.15 ± 2.20 years. Half of the students (50.0%) reported a moderate level of knowledge about artificial intelligence. A very high proportion of students perceived artificial intelligence as useful in daily life, with 53.5% agreeing and 39.1% strongly agreeing. The majority of participants (92.3%) had previously used artificial intelligence technologies. Although only a small proportion (14.1%) had received formal or online training in AI-related topics, most students (63.0%) expressed interest in receiving AI-related education (Table 1).

Table 1. Sociodemographic and descriptive characteristics of participants.

Variables	Number	%
Gender		
Women	216	76.1
Men	66	23.2
Not Declared	2	0.7
Age		
18-38 years (Mean: 21.15 ± 2.2)	284	100
Department of students		
Medicine	6	2.1
Nursing	263	92.6
Engineering	4	1.4
Economics and Administrative Sciences	4	1.4
Humanities and Literature	4	1.4
Law	3	1.1
Self-Rated Knowledge Level About AI		
No Knowledge	7	2.4
Low	58	20.4
Moderate	142	50.0
Good	64	22.5
Very Good	13	4.6
Perception of the Usefulness of Artificial Intelligence in Daily Life		
Strongly disagree	4	1.4
Disagree	2	0.7
Neutral	15	5.3
Agree	152	53.5
Strongly agree	111	39.1
Previous Use of Artificial Intelligence Technologies		
Yes	262	92.3
No	22	7.7
AI-Related Training Experience		
Yes, in formal education	23	8.1
Yes, on online platforms	17	6.0
No, but I am interested	179	63.0
No, I am not interested	65	22.9

3.2. Results of validity analysis

3.2.1. Content validity

Prior to the pilot testing of the scale, a comprehensive content validity assessment was undertaken following the procedures recommended by Polit et al. (2007). Eight experts in the fields of nursing, educational sciences, and health informatics independently evaluated each item for relevance, clarity, and representativeness. The Item-Level Content Validity Index (I-CVI) values ranged from 0.98 to 1.00, demonstrating an excellent level of agreement among the experts. The Scale-Level Content Validity Index (S-CVI/) was computed as 0.99, indicating a high degree of overall content validity and consensus regarding the appropriateness of the scale items.

3.2.2. Construct validity

The construct validity of the Turkish version of the Artificial Intelligence Self-Efficacy Scale (AISES) was examined using Exploratory Factor Analysis (EFA) and Confirmatory Factor Analysis (CFA). Prior to factor analysis, data suitability was assessed using the Kaiser–Meyer–Olkin (KMO) Measure of Sampling Adequacy and Bartlett's Test of Sphericity. The KMO value was 0.919, indicating excellent sampling adequacy, while Bartlett's Test of Sphericity was statistically significant ($\chi^2 = 5161.222$, df = 231, $p < 0.001$), confirming that the correlation matrix was factorable. The EFA identified a four-factor structure: Assistance, Anthropomorphic Interaction, Comfort with AI, and Technological Skills, which together accounted for 73.69% of the total variance. The eigenvalues for these four factors were 9.773, 3.400, 1.813, and 1.224, explaining 44.424%, 15.452%, 8.239%, and 5.570% of the variance, respectively. Factor loadings ranged from 0.640 to 0.936, demonstrating strong associations between items and their respective factors. The first factor, Assistance, consisted of six items that reflected the perceived support of AI technologies in learning and task facilitation. The second factor, Anthropomorphic Interaction, consisted of five items representing the human-like qualities of AI interactions. The third factor, Comfort with AI, consisted of six items that emphasized emotional ease and positive feelings during interactions with AI systems. The fourth factor, Technological Skills, comprises four items assessing confidence in handling AI technologies and minimizing operational concerns (see Table 2).

3.3. Reliability analysis

The internal consistency reliability of the Turkish version of the Artificial Intelligence Self-Efficacy Scale (AISES) was evaluated using Cronbach's alpha coefficients and item–total score correlation analyses. The overall Cronbach's alpha coefficient for the entire scale was 0.937, indicating excellent internal consistency reliability. The subscale reliability coefficients were also high, with values of 0.942 for Assistance, 0.909 for Anthropomorphic Interaction, 0.925 for Comfort with AI, and 0.831 for Technological Skills. Item–total score correlation coefficients ranged from 0.567 to 0.785, demonstrating moderate to strong positive relationships between each item and the total scale score. All items–total correlation values exceeded the recommended threshold of 0.30, suggesting that each item contributes meaningfully to the internal consistency of its respective subscale. These findings confirm that the scale demonstrates robust internal reliability and stable item performance across all dimensions (see Table 3).

3.3.1. Confirmatory factor analysis

Confirmatory Factor Analysis (CFA) was performed to evaluate the construct validity and verify the four-factor structure of the Turkish version of the AISES, which was identified through Exploratory Factor Analysis (EFA). The analysis indicated that the four-factor model demonstrated an acceptable to good fit with the observed data (Figure 1). The chi-square value (χ^2) was 580.876 with 200 degrees of freedom (df), resulting in a χ^2/df ratio of 2.904, which falls within the acceptable range (< 3) for model fit. The Root Mean Square Error of Approximation (RMSEA) was 0.082, suggesting a reasonable level of model fit. Additional fit indices further supported the adequacy of the four-factor model: Goodness of Fit Index (GFI) = 0.846, Relative Fit Index (RFI) = 0.870, Incremental Fit Index (IFI) = 0.926, Comparative Fit Index (CFI) = 0.925, Normed Fit Index (NFI) = 0.891, and Tucker–Lewis Index (TLI)

Table 2. Results of the exploratory factor analysis (n = 142).

	Items	Assistance	Anthropomorphic interaction	Comfort with AI	Technological skills
AS 1	Some AI technologies/products make learning easier. Bazı yapay zekâ teknolojileri/araçları öğrenmeyi kolaylaştırır.	0.936			
AS 2	I find that AI technologies/products are helpful for learning. Yapay zekâ teknolojilerinin/araçlarını öğrenme açısından faydalı buluyorum.	0.911			
AS 3	AI technologies/products are good aids to learning. Yapay zekâ teknolojileri/araçları, öğrenme sürecinde etkili bir yardımcıdır.	0.900			
AS 4	Using AI technologies/products makes learning more interesting. Yapay zekâ teknolojilerini/araçlarını kullanmak öğrenmeyi daha ilgi çekici hale getirir.	0.844			
AS 5	I'm confident in my ability to learn simple programming of AI technologies/products if I were provided the necessary training. Gerekli eğitim verildiği takdirde, yapay zekâ teknolojilerinin/araçlarının basit programlamasını öğrenebileceğime inanıyorum.	0.804			
AS 6	AI technologies/products help me to save a lot of time. Yapay zekâ teknolojileri/araçları, benim çok fazla zaman kazanmama yardımcı olur.	0.759			
AS 7	I find it easy to get AI technologies/products to do what I want it to do. Yapay zekâ teknolojilerinin/araçlarının istediğim şeyi yapmasını sağlamanın kolay olduğunu düşünüyorum.	0.752			
AI 1	I think the interactive process of AI technologies/products is very vivid, just like chatting with a real person. Yapay zekâ teknolojilerinin/araçlarının etkileşim sürecinin, gerçek bir kişiyle sohbet etmek gibi gerçekçi olduğunu düşünüyorum.		0.932		
AI 2	I think the way that AI technologies/products express content when interacting is unique, just like a real person. Yapay zekâ teknolojilerinin/araçlarının içerik ifade etme şeklinin, gerçek bir kişide olduğu gibi benzersiz olduğunu düşünüyorum.		0.813		
AI 3	I think there is no difference between the dialogue method of AI technologies/products compared with the dialogue with real people. Yapay zekâ teknolojileriyle kurulan diyalogların, gerçek kişilerle yapılan sohbetlerden farklı olmadığını düşünüyorum.		0.807		
AI 4	I think the tone of AI technologies/products when interacting is the same as that of real people. Yapay zekâ teknolojilerinin etkileşim sırasında kullandığı üslubun, gerçek insanların iletişim tarzına oldukça benzediğini düşünüyorum		0.784		
AI 5	I feel that the way of expression of AI technologies/products in the interactive text is the same as that of real people. Yapay zekâ teknolojilerinin etkileşimli metinlerde kullandığı ifade tarzının, gerçek insanlarınkiyle örtüştüğünü düşünüyorum.		0.736		
CF 1	When interacting with AI technologies/products, I feel very calm. Yapay zekâ teknolojileri/araçları ile etkileşimde bulunduğumda kendimi çok sakin hissederim.			0.912	
CF 2	When interacting with AI technologies/products, I find it easy. Yapay zekâ teknolojileri/araçları ile etkileşim kurmayı kolay buluyorum.			0.892	
CF 3	When interacting with AI technologies/products, I feel comfortable in my heart. Yapay zekâ teknolojileri/araçları ile etkileşim kurarken içten bir rahatlık hissederim.			0.882	
CF 4	When interacting with AI technologies/products, I feel very peaceful. Yapay zekâ teknolojileri/araçları ile etkileşim kurarken kendimi huzurlu hissederim.			0.694	

(*continued*)

Table 2. Continued.

	Items	Assistance	Anthropomorphic interaction	Comfort with AI	Technological skills
CF5	When interacting with AI technologies/products, I feel very relaxed. Yapay zekâ teknolojileri/araçları etkileşim kurarken kendimi rahatlamış hissederim.			0.645	
CF 6	I can happily interact with AI technologies/products smoothly. Yapay zekâ teknolojileri/araçları ile sorunsuz ve mutlu bir şekilde etkileşim kurabilirim.			0.640	
TS1	When using AI technologies/products, I am not worried that I might press the wrong button and cause risks. Yapay zekâ teknolojilerini/araçlarını kullanırken yanlış bir tuşa basıp risk oluşturacağımdan endişe duymam.				0.851
TS2	When using AI technologies/products, I am not worried that I might press the wrong button and damage it. Yapay zekâ teknolojilerini/araçlarını kullanırken yanlış bir tuşa basıp onu bozacağımdan endişe duymam.				0.840
TS3	When using an AI technology/product, there is nothing that I do not know why. Bir yapay zekâ teknolojisi/araçlarını kullanırken nedenini bilmediğim hiçbir şey olmaz.				0.688
TS4	AI technologies/products jargon does not baffle me. Yapay zekâ teknolojilerine/araçlarına özgü terimler beni şaşırtmaz.				0.688
Eigenvalues		9.773	3.400	1.813	1.224
Explained variance (%)		44.424%	15.452%	8.239	5.570%
KMO = 0.919, Bartlett's Test of Sphericity = 5161.222, df= 231, $p < 0.001$					

KMO: Kaiser-Meyer Olkin coefficient.

Table 3. Results of the reliability analyses of the scale and correlations of the item total score (n = 142).

Items	Cronbach α	Mean ± SD	Item-total score correlation
AS 1	Assistance = 0.942	5.61 ± 1.32	0.623
AS 2		5.63 ± 1.33	0.619
AS 3		5.61 ± 1.28	0.641
AS 4		5.42 ± 1.40	0.656
AS 5		5.49 ± 1.35	0.602
AS 6		5.66 ± 1.35	0.622
AS 7		5.43 ± 1.38	0.629
AI 1	Anthropomorphic Interaction = 0.909	4.07 ± 1.80	0.638
AI 2		3.85 ± 1.85	0.682
AI 3		3.55 ± 1.86	0.612
AI 4		3.94 ± 1.74	0.647
AI 5		4.00 ± 1.66	0.686
CF 1	Comfort with AI = 0.925	4.33 ± 1.67	0.758
CF 2		5.18 ± 1.30	0.705
CF 3		4.35 ± 1.60	0.770
CF 4		4.28 ± 1.56	0.782
CF5		4.27 ± 1.61	0.785
CF 6		4.46 ± 1.59	0.762
TS1	Technological Skills = 0.831	4.54 ± 1.83	0.578
TS2		4.66 ± 1.80	0.580
TS3		3.84 ± 1.69	0.597
TS4		4.37 ± 1.55	0.567
Total Cronbach Alpha (α) Value = 0.937			

= 0.913. Although GFI and NFI were slightly below the ideal cutoff of 0.90, all other indices exceeded or approached acceptable thresholds, indicating a satisfactory overall model fit. These results confirm that the four-factor model provides a theoretically coherent and empirically supported structure, validating the multidimensional construct of the Turkish version of the AISES (see Table 4).

A split-half reliability analysis further supported the internal consistency of the scale and its subdimensions. The total Cronbach's alpha coefficient for the scale was 0.934, indicating excellent internal consistency. The subdimensions also demonstrate high reliability, with Cronbach's alpha values of 0.937 for the first factor, 0.882 for the second factor, and 0.888 for the third factor. A split-half reliability analysis

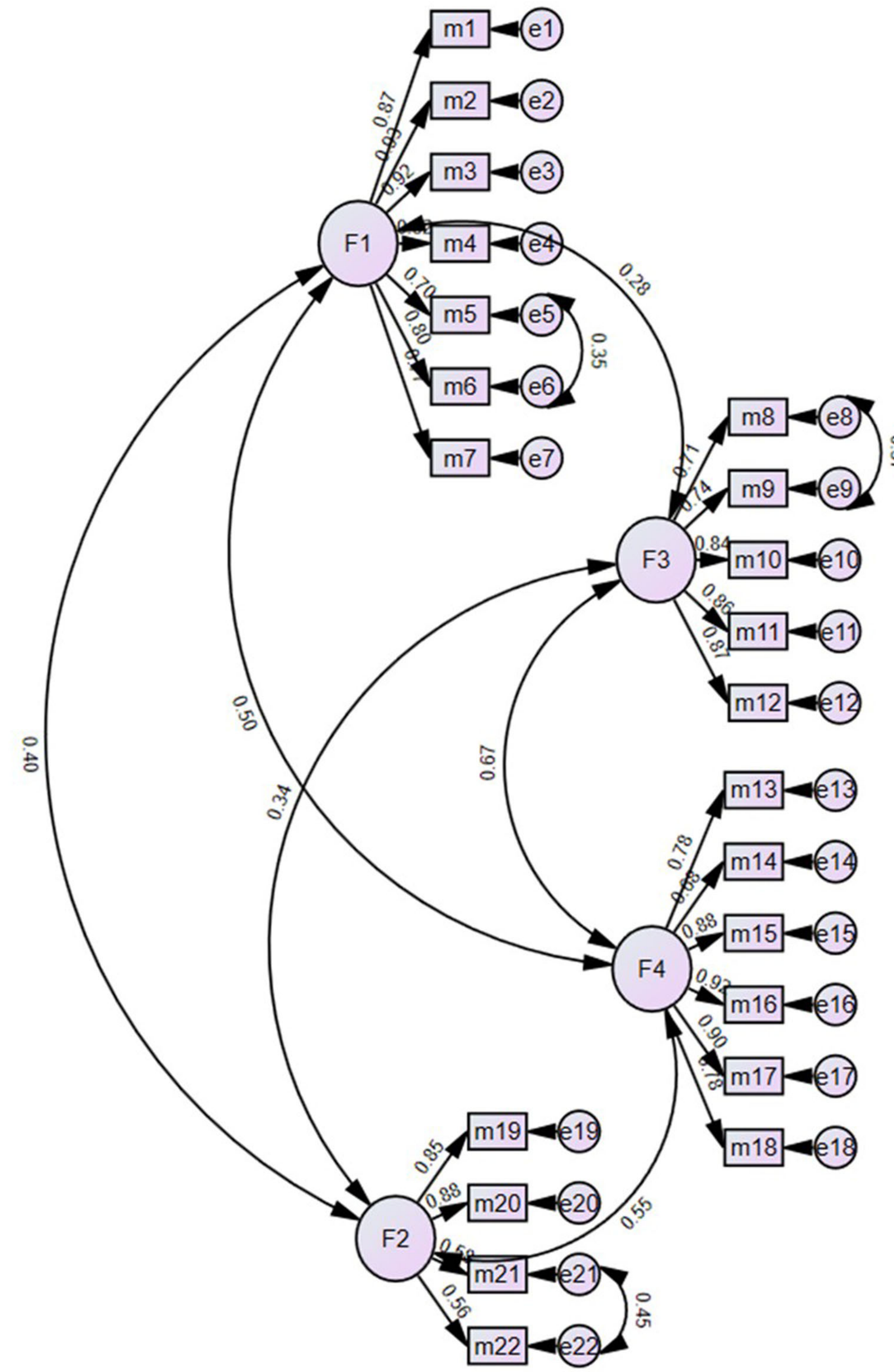

Figure 1. Path diagram and factor loads of the four-factor models of the scale.

was performed for the first factor, which had a first-half Cronbach's alpha of 0.900 and a second-half alpha of 0.912, suggesting strong internal consistency within each split. The Spearman-Brown coefficient was 0.814, and the Guttman split-half coefficient was 0.809, confirming the scale's reliability. Additionally, the correlation between the two halves was 0.686, reflecting moderate to strong agreement between the item halves (see Table 5).

Table 4. Model fit indices of the scale (n = 142).

	X^2	DF[a]	X^2/DF	RMSEA[b]	GFI[c]	RFI[d]	IFI[e]	CFI[f]	NFI[g]	TLI[h]
Four-factor model	580.876	200	2.904	0.082	0.846	0.87	0.926	0.925	0.891	0.913

[a]Degree of Free.
[b]Root Mean Square Error of Approximation.
[c]Goodness of Fit Index.
[d]Relative Fit Index.
[e]Incremental Fit Index.
[f]The comparative fit index.
[g]Normed Fit Index.
[h]TLI (NFI): Tucker-Lewis Index.

Table 5. Results of reliability analysis of scale and subdimensions (n = 284).

		Split into two halves				
Subdimensions	Cronbach α	First half cronbach α	Second half cronbach α	Spearman-brown	Guttman split-half	Correlation between two halves
Total scale	0.937	0.899	0.913	0.827	0.823	0.705

Hotelling's T^2 test was conducted to assess the presence of potential response bias and to evaluate the overall differentiation among item responses. The analysis produced a Hotelling's T^2 value of 458.990, corresponding to an F-value of 20.312 with $df_1 = 21$ and $df_2 = 263$, which was statistically significant ($p < 0.001$). These results indicate significant multivariate differences across item responses, demonstrating that participants responded distinctively to each item rather than following a uniform response pattern.

4. Discussion

This study aimed to evaluate the validity and reliability of the Turkish version of the AISES among university students. Overall, the results demonstrated that the Turkish version of the scale is a psychometrically sound instrument for assessing students' confidence in using and interacting with AI technologies in educational settings.

To assess content validity, the I-CVI and S-CVI were calculated based on the evaluations of eight experts in education, measurement and evaluation, and artificial intelligence applications. Experts assessed each item for relevance, clarity, and cultural appropriateness. As a result of their evaluations, minor revisions were made to wording and semantics to improve conceptual clarity. Expert review yielded exceptionally high content validity indices (I-CVI = 0.98–1.00; S-CVI = 0.99), far exceeding the minimum recommended threshold of 0.80 (Polit et al., 2007). When compared to the scale adaptation study by Aca et al. (2025), which found an I-CVI of 0.94 and an S-CVI of 0.93, indicating excellent content validity.

The high CVI values obtained in this study confirm that the adapted items accurately represent the construct of AI self-efficacy and are appropriate for use among students. These findings provide strong evidence for the scale's conceptual adequacy and a solid foundation for further testing of construct validity.

The four-factor structure identified through Exploratory Factor Analysis, which included Assistance, Anthropomorphic Interaction, Comfort with AI, and Technological Skills, demonstrated both theoretical coherence and empirical strength, explaining 73.7% of the total variance. Before the factor analyses, the data's suitability was tested using the Kaiser–Meyer–Olkin (KMO) and Bartlett's test of sphericity. The KMO value of 0.921 indicated excellent sampling adequacy, and Bartlett's test yielded a statistically significant result ($p < 0.001$), confirming the appropriateness of the data for factor analysis (Yaşlıoğlu, 2017). Factor loadings between 0.64 and 0.94 indicate that all items contributed meaningfully to their respective constructions. The EFA revealed a four-factor structure Assistance, Anthropomorphic Interaction, Comfort with AI, and Technological Skills which explained 73.7% of the total variance. Factor loadings ranged between 0.64 and 0.94, demonstrating that all items contributed meaningfully to their respective constructs. The factor configuration aligns with Wang and Chuang (2024) original model and with cultural adaptations, confirming that the multidimensional nature of AI self-efficacy is stable across contexts.

Confirmatory Factor Analysis supported the four-factor model, with fit indices indicating an acceptable to good model fit (χ^2/df = 2.90, RMSEA = 0.082, CFI = 0.925, TLI = 0.913). Although the GFI (0.846) and NFI (0.891) values were slightly below ideal thresholds, these outcomes are typical for multidimensional psychological scales with numerous items. These findings indicate an acceptable to good model fit (Hu & Bentler, 1999; Hair et al., 2020). These results are comparable to those reported by Wang and Chuang (2024) (χ^2/df = 1.984, RMSEA = 0.079, CFI = 0.941, TLI = 0.930, SRMR = 0.071), who also found strong construct reliability (CR > 0.80) and satisfactory convergent and discriminant validity (AVE > 0.50). Likewise, Aca et al. (2025) reported slightly stronger fit indices (χ^2/df = 2.284, RMSEA = 0.066, SRMR = 0.041, CFI = 0.958, TLI = 0.948, GFI = 0.903, NFI = 0.927, RFI = 0.912), suggesting good model performance in their sample. The differences between studies may be attributed to variations in sample composition, number of items, and cultural context. Despite minor variations in fit statistics, the Turkish version developed in this study retained the theoretical and empirical integrity of the original construct, confirming the stability of the four-factor structure across cultural contexts. These findings demonstrate that the adapted scale effectively captures the multidimensional nature of AI self-efficacy among university students.

The reliability of the Turkish version of AISES was examined using item–total correlation and Cronbach's alpha tests to determine internal consistency. The total Cronbach's α value of 0.937, with subscale coefficients ranging from 0.831 to 0.942, indicates excellent internal consistency and exceeds the commonly accepted threshold of 0.70 (Nunnally & Bernstein, 1994). These results are consistent with those reported in the original study by Wang and Chuang (2024), who found an overall Cronbach's α of 0.958 and subscale coefficients of 0.942 for Assistance, 0.970 for Anthropomorphic Interaction, 0.963 for Comfort with AI, and 0.869 for Technological Skills. Similarly, the Turkish adaptation by Aca et al. (2025) reported comparable reliability coefficients, with subscale alphas of 0.94, 0.97, 0.96, and 0.86, respectively, and an overall reliability of 0.925 after item refinement. The item–total correlations in the present study ranged from 0.567 to 0.785, demonstrating that each item contributed meaningfully to the overall construct and was strongly correlated with the total scale score (Polit & Beck, 2014). In addition, split-half reliability analysis supported the stability of the scale, yielding Spearman–Brown (0.814) and Guttman (0.809) coefficients, both of which indicate high internal reliability.

The reliability coefficients observed in this study align with the "excellent" classification ($\geq$0.90) proposed by Tavakol and Dennick (2011). The results also confirm that the scale's subdimensions maintain consistent psychometric strength across cultural contexts, as previously demonstrated in the original and other adapted versions. Furthermore, Hotelling's T^2 test showed statistically significant differentiation among item responses ($p < 0.001$), suggesting that participants engaged thoughtfully with each item rather than responding uniformly. This finding supports the discriminative capacity of the scale items and reflects the multidimensional nature of AI self-efficacy, consistent with previous psychometric findings (Chiu et al., 2025; Grassini, 2023). Collectively, these results demonstrate that the Turkish version of the AISES is a highly reliable and internally consistent instrument, capable of accurately measuring students' perceived competence and confidence in using and adapting to AI technologies.

5. Limitations

Several limitations should be noted. The sample was predominantly female and drawn mainly from nursing, which may limit the generalizability of findings to other academic disciplines. Future studies should include more balanced samples and examine measurement invariance across gender, discipline, and level of AI exposure. Additionally, this study employed a cross-sectional design; longitudinal validation would further verify the scale's stability over time. Finally, the study relied on self-reported data, which may be influenced by social desirability bias.

6. Conclusion

This validated instrument can serve as a valuable tool for educators and researchers aiming to evaluate and strengthen students' readiness for AI-enhanced learning environments in Türkiye and beyond. The clear presence of emotional and anthropomorphic factors suggests that AI education should address not only technical competence but also students' affective responses and trust in AI systems. Third, given that most participants reported only moderate AI knowledge but high willingness to learn, integrating structured AI literacy modules into nursing and other health curricula could enhance both competence and confidence. This study expands the evidence base for cross-cultural validation of AI self-efficacy measures. Replication in student populations provides essential evidence for scale generalizability, complementing earlier Turkish adaptations conducted in professional or mixed adult samples. Moreover, the strong psychometric outcomes confirm that the AISES can be used to monitor changes in students' self-efficacy over time or to evaluate educational interventions involving AI technologies.

Acknowledgments

We thank all the study participants for contributing to this study.

Ethical considerations

Ethical approval for this study was obtained from the Institutional Review Board of Koç University.

Author contributions

CRediT: **Seda Güney**: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Project administration, Writing – original draft, Writing – review & editing; **Remziye Semerci Şahin**: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Project administration, Validation, Writing – original draft, Writing – review & editing.

Declaration of AI use

AI-assisted language editing was used for minor improvements in readability and grammar using ChatGPT (OpenAI). All content, interpretations, and conclusions are the responsibility of the authors.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

No funding was available for this study.

ORCID

Seda Güney http://orcid.org/0000-0003-3704-0498
Remziye Semerci Şahin http://orcid.org/0000-0003-1999-9179

Data availability

All data supporting the findings of this study are available within the article and its Supplementary Information files. Any additional data or materials that support the results of this study are available from the corresponding author upon reasonable request.

References

Aca, Z., Solmaz, U., & Koçak, O. (2025). Turkish adaptation of the AI self-efficacy scale: A psychometric evaluation. *International Journal of Human–Computer Interaction*, 1–17. https://doi.org/10.1080/10447318.2025.2558027

Albino, M. G., Albino, F. S., Asio, J. M. R., & Gadia, E. D. (2025). Influence of AI anxiety on AI self-efficacy among college students. *International Journal of Technology in Education*, *8*(2), 557–573. https://doi.org/10.46328/ijte.1109

Asio, J. M. R., & Sardina, D. P. (2025). Gender differences on the impact of AI self-efficacy on AI anxiety through AI self-competency: A moderated mediation analysis. *Journal of Pedagogical Research*, *9*(2), 55–71. https://doi.org/10.33902/JPR.202533231

Asio, J. M. R. (2025). Navigating AI self-efficacy: Mediating student attitudes and AI literacy. *Journal of Pedagogical Sociology and Psychology*, *7*(3), 49–63. https://doi.org/10.33902/JPSP.202535272

Atik, S., & Erdemir, N. (2025). Validity and reliability analysis of the Turkish life satisfaction scale developed through artificial intelligence. *BMC Psychology*, *13*(1), 1139. https://doi.org/10.1186/s40359-025-03246-2

Bećirović, S., Polz, E., & Tinkel, I. (2025). Exploring students' AI literacy and its effects on their AI output quality, self-efficacy, and academic performance. *Smart Learning Environments*, *12*(1), 29. https://doi.org/10.1186/s40561-025-00384-3

Çapık, C., Gözüm, S., & Aksayan, S. (2018). Intercultural scale adaptation stages, language and culture adaptation: Updated guideline. *Florence Nightingale Journal of Nursing*, *26*(3), 199–210. https://doi.org/10.26650/FNJN397481

Chen, G., Gully, S. M., & Eden, D. (2001). Validation of a new general self-efficacy scale. *Organizational Research Methods*, *4*(1), 62–83. https://doi.org/10.1177/109442810141004

Chiu, T. K. F., Ahmad, Z., & Çoban, M. (2025). Development and validation of teacher artificial intelligence (AI) competence self-efficacy (TAICS) scale. *Education and Information Technologies*, *30*(5), 6667–6685. https://doi.org/10.1007/s10639-024-13094-z

Chu, H.-C., Hwang, G.-H., Tu, Y.-F., & Yang, K.-H. (2022). Roles and research trends of artificial intelligence in higher education: A systematic review of the top 50 most-cited articles. *Australasian Journal of Educational Technology*, *38*(3), 22–42. https://doi.org/10.14742/ajet.7526

Costello, A. B., & Osborne, J. (2005). Best practices in exploratory factor analysis: Four recommendations for getting the most from your analysis. *Practical Assessment, Research & Evaluation*, *10*(1), 1–7. https://doi.org/10.7275/jyj1-4868

Eisinga, R., Grotenhuis, M. T., & Pelzer, B. (2013). The reliability of a two-item scale: Pearson, Cronbach, or Spearman–Brown? *International Journal of Public Health*, *58*(4), 637–642. https://doi.org/10.1007/s00038-012-0416-3

Eminoğlu, A., & Çelikkanat, Ş. (2024). Assessment of the relationship between executive nurses' leadership self-efficacy and medical artificial intelligence readiness. *International Journal of Medical Informatics*, *184*(2024), 105386. https://doi.org/10.1016/j.ijmedinf.2024.105386

Grassini, S. (2023). Development and validation of the AI attitude scale (AIAS-4): A brief measure of general attitude toward artificial intelligence. *Frontiers in Psychology*, *14*(2023), 1191628. https://doi.org/10.3389/fpsyg.2023.1191628

Gunawan, J., Marzilli, C., & Aungsuroch, Y. (2021). Establishing an appropriate sample size for developing and validating a questionnaire in nursing research. *Belitung Nursing Journal*, *7*(5), 356–360. https://doi.org/10.33546/bnj.1927

Falebita, O. S., & Kok, P. J. (2025). Artificial intelligence tools usage: A structural equation modeling of undergraduates' technological readiness, self-efficacy, and attitudes. *Journal for STEM Education Research*, *8*(2), 257–282. https://doi.org/10.1007/s41979-024-00132-1

Hair, J. F., Black, B., Babin, B. J., & Anderson, R. E. (2010). *Multivariate data analysis* (7th ed.). Pearson.

Hair, J. F., Jr., Howard, M. C., & Nitzl, C. (2020). Assessing measurement model quality in PLS-SEM using confirmatory composite analysis. *Journal of Business Research*, *109*(2020), 101–110. https://doi.org/10.1016/j.jbusres.2019.11.069

Hajjar, S. T. (2018). Statistical analysis: Internal-consistency reliability and construct validity. *International Journal of Quantitative and Qualitative Research Methods*, *6*(1), 46–57.

Howard, M. C. (2016). A review of exploratory factor analysis decisions and overview of current practices: What we are doing and how can we improve? *International Journal of Human-Computer Interaction*, *32*(1), 51–62. https://doi.org/10.1080/10447318.2015.1087664

Hu, L., & Bentler, P. M. (1999). Cut-off criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, *6*(1), 1–55. https://doi.org/10.1080/10705519909540118

Luo, Q. Z., & Hsiao-Chin, L. Y. (2023). The influence of AI-powered adaptive learning platforms on student performance in Chinese classrooms. *Journal of Education*, *6*(3), 1–12. https://doi.org/10.53819/81018102t4181

Morales-García, W. C., Sairitupa-Sanchez, L. Z., Morales-García, S. B., & Morales-García, M. (2024). Adaptation and psychometric properties of a brief version of the general self-efficacy scale for use with artificial intelligence (GSE-6AI) among university students. *Frontiers in Education*, *9*(2024), 1293437. https://doi.org/10.3389/feduc.2024.1293437

Nunnally, J. C., & Bernstein, I. H. (1994). The assessment of reliability. *Psychometric Theory*, *3*(1994), 248–292.

Owan, V. J., Abang, K. B., Idika, D. O., Etta, E. O., & Bassey, B. A. (2023). Exploring the potential of artificial intelligence tools in educational measurement and assessment. *Eurasia Journal of Mathematics, Science and Technology Education*, *19*(8), em2307. https://doi.org/10.29333/ejmste/13428

Polit, D. F., & Beck, C. T. (2014). *Essentials of nursing research: Appraising evidence for nursing practice.* (8th ed.). Wolters Kluwer Health/Lippincott Williams & Wilkins.

Polit, D. F., Beck, C. T., & Owen, S. V. (2007). Is the CVI an acceptable indicator of content validity? Appraisal and recommendations. *Research in Nursing & Health*, *30*(4), 459–467. https://doi.org/10.1002/nur.20199

Şencan, H. (2005). *Sosyal ve davranışsal ölçümlerde güvenilirlik ve geçerlilik*. Seçkin Yayıncılık.

Stracke, C. M., Griffiths, D., Pappa, D., Bećirović, S., Polz, E., Perla, L., Grassi, A. D., Massaro, S., Skenduli, M. P., Burgos, D., Punzo, V., Amram, D., Ziouvelou, X., Katsamori, D., Gabriel, S., Nahar, N., Schleiss, J., & Hollins, P. (2025). Analysis of artificial intelligence policies for higher education in Europe. *International Journal of Interactive Multimedia and Artificial Intelligence*, *9*(2), 124–137. https://doi.org/10.9781/ijimai.2025.02.011

Tavakol, M., & Dennick, R. (2011). Making sense of Cronbach's alpha. *International Journal of Medical Education*, *2*(2011), 53–55. https://doi.org/10.5116/ijme.4dfb.8dfd

Terzi, R. (2020). An adaptation of artificial intelligence anxiety scale into Turkish: Reliability and validity study. *International Online Journal of Education and Teaching (IOJET)*, *7*(4), 1501–1515. http://iojet.org/index.php/IOJET/article/view/1031

Uluğ, E., Öner, K., Arslantaş, S., & Harmancı, S. T. (2025). Adaptation of the artificial intelligence literacy scale into Turkish: A cross-sectional application among healthcare workers, students, and children. *Education and Information Technologies*, *30*(15), 21041–21077. https://doi.org/10.1007/s10639-025-13597-3

Uyan, U., & Gültekin, S. U. (2024). Yapay zeka öz-yeterlilik ölçeğinin Türkçe'ye uyarlanması: Geçerlilik ve güvenirlik çalışması. *Journal of Research in Business*, *9*(1), 135–148. https://doi.org/10.54452/jrb.1415212

Von Elm, E., Altman, D. G., Egger, M., Pocock, S. J., Gøtzsche, P. C., & Vandenbroucke, J. P., STROBE Initiative (2007). The STROBE statement: Guidelines for reporting observational studies. *Lancet (London, England)*, *370*(9596), 1453–1457. https://doi.org/10.1016/S0140-6736(07)61602-X

Wang, Y. Y., & Chuang, Y. W. (2024). Artificial intelligence self-efficacy: Scale development and validation. *Education and Information Technologies*, *29*(4), 4785–4808. https://doi.org/10.1007/s10639-023-12015-w

Wang, Y., Wu, H., & Wang, Y. (2024). Engagement and willingness to communicate in the L2 classroom: Identifying the latent profiles and their relationships with achievement emotions. *Journal of Multilingual and Multicultural Development*, *46*(8), 2175–2191. https://doi.org/10.1080/01434632.2024.2379534

World Health Organization. (2017). *Process of translation and adaptation of instruments*. http://www.who.int/substance_abuse/research_tools/translation/en/

Yaşlıoğlu, M. M. (2017). Factor analysis and validity in social sciences: The use of exploratory and confirmatory factor analyses. *Istanbul University Journal of the School of Business*, *46*(November 2017), 74–85. https://dergipark.org.tr/en/pub/iuisletme/article/357061

Yasir, A. S. M. (2016). Cross-cultural adaptation & psychometric validation of instruments: Step-wise description. *International Journal of Psychiatry*, *1*(1), 1–4. https://doi.org/10.33140/ijp/01/01/00001

Yang, J., Chen, Y., & Wang, Y. (2025). Exploring the interplay of motivation, engagement and critical thinking among EFL learners: Evidence from structural equation modelling. *European Journal of Education*, *60*(3), e70187. https://doi.org/10.1111/ejed.70187

Yılmaz, F. G. K., Yılmaz, R., & Ceylan, M. (2023). Generative artificial intelligence acceptance scale: A validity and reliability study. *International Journal of Human–Computer Interaction*, *40*(24), 8703–8715. https://doi.org/10.1080/10447318.2023.2288730

About the authors

Seda Güney, PhD is an Assistant professor at Koç University School of Nursing in Türkiye. Her research focuses on public health and gerontological nursing, with particular interest in dementia care, aging, caregiver support, climate change and health, scale adaptation, and nursing education.

Remziye Semerci Şahin, PhD is an Associate professor of nursing at Koç University School of Nursing in Türkiye. Her research interests include pediatric nursing, pediatric oncology care, symptom management, pain assessment, scale adaptation, and supportive and family-centered nursing interventions.