European Journal of Pediatrics (2025) 184:585
https://doi.org/10.1007/s00431-025-06435-8

RESEARCH

Development of the generative artificial intelligence awareness scale for secondary school students in Türkiye

Remziye Semerci Şahin[1] · Özkan Özbay[2] · Sevil Çınar Özbay[3] · Selma Durmuş Sarıkahya[4]

Received: 12 July 2025 / Revised: 7 August 2025 / Accepted: 22 August 2025

Abstract
The widespread adoption of generative artificial intelligence (AI) in education and daily life necessitates a deeper understanding of students' awareness and attitudes. However, there is a lack of appropriate and psychometrically validated tools to assess these constructs reliably. This study aimed to examine the validity and reliability of a newly developed scale measuring secondary school students' awareness of generative artificial intelligence in the Turkish context. This methodological study employed a cross-sectional, descriptive, and correlational design. The research was conducted with 444 secondary school students aged 14–18. Data were collected using a Demographic Information Form and the Generative Artificial Intelligence Awareness Scale. The study followed a three-stage process in line with the STROBE checklist: (1) item development, (2) expert evaluation for content validity, and (3) construct validity and reliability testing. Construct validity was assessed using Exploratory and Confirmatory Factor Analyses. Reliability was examined through Cronbach's alpha, corrected item-total correlations, and split-half reliability analysis. EFA revealed a three-factor structure: Basic Knowledge of Generative Artificial Intelligence, Positive Attitudes Toward Generative Artificial Intelligence, and Concerns About the Impacts of Generative Artificial Intelligence, explaining 54.29% of the total variance. CFA supported the model, yielding acceptable fit indices (χ^2/df = 3.792, RMSEA = 0.056, GFI = 0.905, CFI = 0.927, TLI = 0.919), indicating good model fit. The Cronbach's alpha for the total scale was 0.946, with subscale coefficients ranging from 0.822 to 0.925. Corrected item-total correlations ranged from 0.512 to 0.733. Split-half reliability for the first factor showed strong consistency (Spearman-Brown = 0.891, Guttman = 0.887, correlation between halves = 0.804). The findings demonstrate that the Turkish version of the generative artificial intelligenceawareness scale is a valid and reliable instrument for assessing students' knowledge, attitudes, and concerns about generative AI.

What is Known:
- *Generative artificial intelligence (GAI) technologies are increasingly used by adolescents in educational settings, yet their awareness, ethical understanding, and usage patterns remain underexplored.*
- *Existing measurement tools for AI literacy often target higher education students and lack specificity in assessing secondary school students' awareness of generative AI.*

What is New:
- *This study developed and psychometrically validated the first comprehensive scale to assess secondary school students' awareness of generative AI, including knowledge, attitudes, and ethical concerns.*
- *The Generative Artificial Intelligence Awareness Scale (GAIAS) provides educators and policymakers with a reliable tool to evaluate and support adolescents' digital and ethical preparedness in the age of generative AI.*

Communicated by Peter de Winter

✉ Sevil Çınar Özbay
cinarsevil87@gmail.com

Remziye Semerci Şahin
remziyesemerci@gmail.com

Özkan Özbay
ozkanozbay87@gmail.com

Selma Durmuş Sarıkahya
durmussel@gmail.com

[1] Child and Disease Nursing Department, Nursing Faculty, Koç University, Istanbul, Türkiye

[2] Distance Education Application and Research Center, Artvin Çoruh University, Artvin, Türkiye

[3] Faculty of Health Sciences, Artvin Çoruh University, Artvin, Türkiye

[4] Department of Public Health Nursing, Faculty of Health Sciences, Artvin Çoruh University, Artvin, Türkiye

Published online: 30 August 2025

Keywords Generative artificial intelligence · Secondary school students · Awareness · Validity · Reliability · Scale development

Introduction

Artificial intelligence (AI) technologies have led to profound transformations in recent years across various domains, including education, healthcare, economics, and social life. These technologies have reshaped how individuals access information, engage in learning processes, and interact socially [1, 2]. One of the most striking aspects of this transformation is the widespread adoption of generative artificial intelligence (GAI) systems, which operate based on advanced algorithms such as large language models (LLMs) [3, 4]. Tools like ChatGPT, DALL·E, and Copilot are leading innovative applications in education by supporting personalized learning, enhancing creative thinking, and facilitating content generation [5, 6].

Building upon these capabilities, GAI distinguishes itself through its ability to analyze student data, design personalized learning pathways, deliver real-time feedback, and provide cognitive support [7, 8]. However, despite these advantages, it also raises significant concerns regarding ethical use, information reliability, privacy, and academic integrity [9, 10]. Therefore, developing a comprehensive awareness of generative AI one that extends beyond technical proficiency to encompass ethical sensitivity, critical thinking, and digital consciousness has become increasingly essential [11, 12].

A clear conceptual distinction between conventional artificial intelligence, generative artificial intelligence (GAI), and awareness of GAI is fundamental to establishing such awareness. Conventional AI refers to systems designed for specific, often rule-based tasks such as classification or prediction, producing only predefined outputs [13–15]. In contrast, GAI exemplified by tools such as ChatGPT and DALL·E can generate novel outputs (e.g., text, images, audio, or code) from learned data, introducing new dimensions of creativity, personalization, and user interaction. These distinctive capabilities also present unique pedagogical, ethical, and cognitive challenges not typically encountered with conventional AI [16, 17].

Importantly, GAI awareness should not be conflated with the technology itself. It refers to users' understanding, perceptions, and ability to engage responsibly with generative systems, including their knowledge of functionalities, ethical implications, and broader societal impact. While GAI is a technological capability, GAI awareness represents a human-centered psychological construct [18]. Conventional AI and GAI differ mainly in their ability to generate new content, with GAI offering broader creative and practical applications. Consequently, a dedicated measurement tool is needed to address GAI's unique impacts. This distinction is essential for both research and practice. The development of tools such as the Generative Artificial Intelligence Awareness Scale (GAIAS) is therefore warranted, as existing AI literacy instruments fail to capture the multifaceted nature of GAI particularly its ethical, creative, and user experience dimensions [19, 20].

In this context, awareness of generative AI can be defined as a multidimensional construct encompassing individuals' levels of knowledge about these technologies, usage habits, understanding of ethical and social responsibility, and perceptions of their future impact [13, 14]. Importantly, young individuals'interactions with these technologies directly influence their ethical decision-making, self-regulation, and critical thinking abilities [15, 21]. While the academic literature has begun to explore the integration of GAI into higher education, empirical research focusing on secondary school students remains limited [22, 23]. This is especially noteworthy given that adolescence represents a crucial developmental stage in which digital literacy and ethical responsibility begin to take shape.

Indeed, secondary school students are active users of GAI tools such as ChatGPT for preparing assignments, generating creative content, and sharing digital outputs on social media. In doing so, they take on roles not only as consumers but also as content producers [15, 24]. A systematic review by Zhang and Tur (2024) confirms this trend, noting that ChatGPT supports curriculum design, content creation, and personalized learning in K–12 education. At the same time, however, the review also highlights significant ethical and pedagogical challenges that accompany the use of such tools [25]. Complementing these findings, a social media content analysis by Mogavi et al. (2024) revealed that ChatGPT is most commonly used in higher education, K–12 education, and practical skills training with a usage rate of 22.09% reported at the K–12 level [26].

Recent research highlights that secondary school students show considerable variation in their levels of knowledge, ethical sensitivity, and usage patterns related to generative AI tools [27, 28]. While students often favor these tools for their speed and convenience, they may overlook their pedagogical and ethical implications. Ng et al. (2024) emphasized that students' cognitive, affective, behavioral, and ethical awareness of AI develops at different rates and must be supported holistically to foster responsible use [29]. Some students tend to perceive generative AI merely as a functional tool, which may risk diminishing critical thinking abilities over time [30]. Thus, effective integration of generative AI into educational settings requires not only technical competence but also ethical literacy and cognitive awareness [31].

These findings underscore the urgent need for a valid and reliable instrument specifically designed to assess secondary school students'awareness of generative AI. Despite the increasing attention to AI literacy, the literature reveals a noticeable scarcity of tools that address the ethical, pedagogical, and conceptual aspects of generative AI in an integrated manner [32, 33]. For example, while the AI Literacy Concept Inventory by Zhang et al. (2024) identifies students' conceptual misunderstandings about AI, it does not fully meet the need for instruments that target the specific dimensions of generative AI [33].

Furthermore, although several existing scales measure AI literacy or attitudes toward AI, most of them are intended for university students or adult populations, and they often emphasize general AI knowledge rather than the unique characteristics of generative AI, such as ethical concerns, creative use, or personalized learning pathways [34]. Additionally, many of these tools are not developmentally appropriate for adolescents in terms of language, content, or contextual relevance [25]. To date, no psychometrically validated tool has been identified that comprehensively assesses generative AI awareness among adolescents aged 14–18 through a multidimensional framework.

To address this gap, the present study developed the GAIAS, specifically designed to evaluate secondary school students' basic knowledge, usage patterns, ethical and social concerns, and perceptions of generative AI's impact. As the first validated scale of its kind in Türkiye targeting this age group, GAIAS offers a valuable contribution to the literature and provides educators and policymakers with a practical tool to support adolescents' digital literacy, ethical decision-making, and responsible engagement with emerging AI technologies [29, 34].

Methods and materials

Study Design

This methodological study was conducted using a cross-sectional, descriptive, and correlational design (internal analysis of the relationships between subscale scores and the total score) to develop the Awareness of Generative Artificial Intelligence Scale for Secondary School Students and to evaluate its psychometric properties. The study adhered to the EQUATOR guidelines and was reported in line with the STROBE checklist.

Study Sample

This study was conducted with secondary school students aged 14–18 years enrolled in four secondary education institutions located in Türkiye, between April and June 2025. In scale development studies, it is generally recommended to recruit at least five to ten participants per item to ensure robust factor analysis [35]. Sample size classifications suggest that fewer than 100 participants are insufficient, 100–200 are sufficient, 200–500 are good, 500–1000 are very good, and over 1000 are considered excellent [36]. Accordingly, we aim to reach a very good sample size between 500 and 1000, and this study was conducted with 888 secondary school students who met the inclusion criteria and volunteered to participate. The study included two distinct samples for the EFA ($n = 444$) and CFA ($n = 444$) analyses. For the EFA sample, the mean age was 15.87 years (SD = 1.09), with 208 females (46.8%) and 236 males (53.2%). Regarding grade distribution, 10.4% were in 9th grade, 30.4% in 10th grade, 26.6% in 11th grade, and 26.8% in 12th grade. In the CFA sample, the mean age was slightly higher at 16.01 years (SD = 1.24), with 266 females (59.9%) and 178 males (40.1%). Class distribution was similar: 10.6% in 9th grade, 29.6% in 10th grade, 25.7% in 11th grade, and 23.3% in 12th grade.

A preliminary group of 60 students was randomly invited from the main sample to participate in the test–retest reliability analysis. Participants were informed that the scale would be administered twice, two weeks apart, and those who volunteered for both sessions were included in the test–retest group. The pilot study was conducted with a total of 60 participants. The mean age of the participants was 16.47 years (SD = 1.31), with an age range of 14 to 18 years. In terms of gender distribution, 36 participants (60.0%) were female and 24 (40.0%) were male. These descriptive statistics provided preliminary insights into the appropriateness and clarity of the scale items prior to conducting the main analyses. Class distribution was 20% in 9th grade, 30% in 10th grade, 24% in 11th grade, and 26% in 12th grade.

The inclusion criteria for participation were as follows: (a) being aged between 14 and 18 years, (b) having the ability to comprehend and respond to the questionnaire items, (c) providing written informed consent (and parental consent where applicable), and (d) voluntarily agreeing to participate. Exclusion criteria included (a) refusal to participate, (b) presence of communication barriers, (c) known cognitive impairments, and (d) significant reading or writing difficulties.

To assess test–retest reliability, the scale was re-administered to a subgroup of 60 students with similar characteristics to the target population two weeks after the initial administration.

Data Collection Tools

Personal Information Form

This form was prepared by the researchers to determine the demographic characteristics of students, such as age, gender, and grade level.

Generative Artificial Intelligence Awareness Scale

A comprehensive literature review was conducted to assess awareness of generative artificial intelligence, based on which a practical scale was developed. The scale comprises 27 items. The scale is based on a 5-point Likert scale (1 = Strongly disagree to 5 = Strongly agree). The total score ranges from 27 to 135, with higher scores indicating greater awareness of generative artificial intelligence. Mean scores near 5 reflect high awareness, around 3 suggest moderate awareness, and near 1 indicate low awareness. The total score ranges from 27 to 135, with higher total and subdimension mean scores indicating greater awareness of generative artificial intelligence among students. The scale includes three theoretically grounded subdimensions:

Basic Knowledge of Generative Artificial Intelligence: This subdimension assesses students' cognitive understanding of key concepts, tools, and capabilities related to generative AI, reflecting their informational awareness.
Positive Attitudes Toward Generative Artificial Intelligence: This subdimension captures students' affective responses and perceived benefits of generative AI in educational and societal contexts, including openness, perceived usefulness, and excitement.
Concerns About the Impacts of Generative Artificial Intelligence: This dimension addresses cognitive and emotional apprehensions related to ethical dilemmas, misinformation, dependence, and potential negative effects of AI tools on learning and psychological development.

Data Collection Procedure

Phase 1: Development of Item Pool

According to DeVellis and Thorpe (2021), item development should be grounded in the extent to which each item accurately reflects the construct being measured and should be informed by both theoretical understanding and empirical evidence. In the development of the draft version of the Awareness of Generative Artificial Intelligence Scale for Secondary School Students, an extensive literature review was conducted using international databases including PubMed, Cochrane, CINAHL, Embase, MEDLINE, and PsycINFO. The literature search focused on themes related to generative artificial intelligence, digital literacy, AI in education, student perceptions of emerging technologies, and ethical concerns. Keywords used in the review included "generative artificial intelligence," "AI awareness," "student perception," "technology in education," "AI ethics," "digital competence," and "secondary school students." Informed by the literature and the research team's experience in psychometric scale development and digital education, an initial pool of 34 items was generated. These items were designed to reflect students' basic knowledge of generative AI, their attitudes toward its use, and their concerns about its ethical and societal implications. The items were constructed using a 5-point Likert-type scale ranging from 1 ("strongly disagree") to 5 ("strongly agree"). A 5-point Likert scale was selected based on its balance between measurement sensitivity and respondent ease of use. It is commonly used in educational and psychological assessments and is particularly appropriate for adolescents, offering clarity and reduced cognitive load while allowing sufficient response variability. In designing the items, negatively worded or reverse-coded statements were intentionally avoided. This decision was based on evidence suggesting that such items may cause confusion, increase cognitive load, and lead to inconsistent responses, particularly among adolescent populations. The goal was to ensure linguistic clarity and response accuracy for secondary school students.

Phase 2: Assessing Content Validity

Evaluation by Expert Opinions

In this phase, the initial 34-item draft of the Awareness of Generative Artificial Intelligence Scale for Secondary School Students was evaluated by a panel of 10 experts with academic backgrounds in educational technology, psychometrics, and digital literacy. The experts were asked to assess each item based on four criteria: content relevance, scope, clarity, and appropriateness of wording. Each item was rated using a 4-point scale: 1 = not appropriate, 2 = item needs substantial revision, 3 = appropriate but needs minor changes, and 4 = very appropriate based on the Polit and Beck technique.

Content validity of the scale was evaluated through expert judgment using the Davis method, which involves the calculation of both the Item-Level Content Validity Index (I-CVI) and the Scale-Level Content Validity Index (S-CVI) [37]. According to the literature, an I-CVI value of 0.80 or higher is typically considered sufficient to demonstrate consensus among experts regarding item relevance [31, 33]. In the current study, both the I-CVI and S-CVI were found to be 0.99, reflecting a very high level of expert agreement and indicating that the items were deemed highly appropriate for measuring the intended construct. Based on expert feedback, content revisions were made to several items to enhance clarity and relevance. A total of 7 items were removed due to redundancy or conceptual overlap, resulting in a refined scale of 27 items for the next stage of analysis.

Preliminary Test Phase

In accordance with established recommendations in the literature, a preliminary test was conducted to assess the readability, linguistic clarity, and comprehensibility of the draft scale prior to its implementation in the main study [35, 36]. This stage involved 60 secondary school students who met the inclusion criteria but were not included in the primary study sample to avoid potential bias. Following the expert validation phase, the revised 27-item version of the Awareness of Generative Artificial Intelligence Scale for Secondary School Students was administered face-to-face by the researchers. Participants were invited to evaluate the clarity of item phrasing and the comprehensibility of response categories. Feedback obtained during this stage informed minor revisions to the wording of certain items to enhance interpretability. No additions or deletions were made to the item pool at this stage. The finalized version of the scale, incorporating these minor modifications, was subsequently employed in the main validation study.

Phase 3: Evaluation of the Validity and Reliability of the Scale

Content validity refers to the degree of agreement among experts regarding whether the items adequately represent the construct being measured [35, 37]. Content validity was previously established through expert evaluation and calculated using I-CVI and S-CVI indices. To assess construct validity, both Exploratory Factor Analysis (EFA) and Confirmatory Factor Analysis (CFA) were conducted. Prior to factor analysis, the suitability of the data was confirmed by calculating the Kaiser–Meyer–Olkin (KMO) measure of sampling adequacy and performing Bartlett's Test of Sphericity [34, 35]. For the CFA, multiple fit indices were used to evaluate model adequacy, including Chi-square (χ^2), degrees of freedom (df), the χ^2/df ratio, Root Mean Square Error of Approximation (RMSEA), Comparative Fit Index (CFI), Incremental Fit Index (IFI), Tucker–Lewis Index (TLI), and Goodness of Fit Index (GFI). Threshold values of ≥ 0.90 for CFI, IFI, TLI, and GFI, and ≤ 0.08 for RMSEA were considered acceptable indicators of model fit [38].

The reliability of the scale was assessed through multiple methods. Internal consistency was evaluated using Cronbach's alpha coefficient. Corrected item-total score correlations were also calculated to assess the contribution of each item to the total scale score.

EFA and CFA were conducted using split samples derived from a single dataset. The total sample ($N = 888$) was randomly divided into two independent subsamples ($n = 444$ each), with one used for EFA and the other for CFA. This approach ensured internal cross-validation while maintaining statistical independence between the two analyses. Split-half reliability was assessed by calculating Cronbach's alpha for the first and second halves of the scale, and the consistency between the two halves was examined using the Spearman-Brown and Guttman split-half coefficients [35].

Hotelling's T^2 test was employed to examine response bias, while the temporal stability of the scale was assessed via test–retest reliability. For this, the scale was re-administered to a subsample of 60 participants two weeks after the initial administration. Pearson correlation coefficients were computed for both the test–retest analysis and the item-total score analysis to establish measurement invariance and reliability over time.

Data Collection

The data collection process was conducted through face-to-face survey administration between April and June 2025, during the Spring semester of the 2024–2025 academic year. Data were collected from secondary school students who voluntarily agreed to participate and had obtained written parental consent. A total of 1,050 students were invited, and 888 students completed the questionnaire in full, resulting in a participation rate of 84.6%. At the beginning of the questionnaire, two separate consent items were presented to verify the participation agreement of both the students and their parents. Only those who confirmed consent were permitted to proceed to the main items of the scale. The surveys were administered in classrooms in a collective setting. Prior to distribution, the researcher introduced the purpose of the study, read the instructions aloud, and emphasized the importance of responding carefully and honestly. Students completed the questionnaires individually during class time, and all completed forms were collected directly by the researcher at the end of the session.

Ethical Considerations

Ethical approval for the study was obtained from the Scientific Research and Publication Ethics Committee of Artvin Çoruh University (Approval No: E-18457941–050.99–171,843). Additionally, permission to conduct the research in schools was granted by the Provincial Directorate of National Education (Approval No: MEB.TT.2025.022666.01). The study was conducted in accordance with the ethical principles outlined in the Declaration of Helsinki. Written informed consent was obtained from both the students and their parents. At the beginning of the questionnaire, participants were asked to confirm consent before proceeding. Participation was voluntary, and students had the right to withdraw from the study at any time without penalty. All data were collected anonymously, and no personally identifiable information was recorded. Completed surveys were stored securely on password-protected computers accessible only to the research team. In line with institutional and ethical guidelines, the data will be retained for five years and then permanently deleted.

Data Analysis

All statistical analyses were performed using IBM SPSS Statistics version 28.0 and AMOS version 24.0 software. Descriptive statistics were computed to summarize participants'sociodemographic information, including means and standard deviations. Content validity was assessed by calculating the Item-Level Content Validity Index (I-CVI) and Scale-Level Content Validity Index (S-CVI), following expert evaluations using the Polit and Beck technique. For construct validity, the sample was randomly divided into two independent subsamples. The first subset was subjected to EFA, while the second was used for CFA. EFA was conducted using principal axis factoring with varimax rotation to explore the underlying factor structure of the scale. Varimax rotation was applied under the assumption that the underlying factors are relatively independent. This orthogonal rotation technique was preferred to enhance the interpretability and simplicity of the factor structure. The Varimax method aims to maximize high and low factor loadings, meaning that the correlations between observed variables and extracted factors tend to approach values close to 1 or 0. By minimizing mid-range loadings, the technique improves conceptual clarity by making it easier to determine which items are most strongly associated with which factors. This is particularly useful in early-stage scale development when the goal is to obtain a clear and distinct factor solution. The appropriateness of the data for factor analysis was verified using the KMO measure of sampling adequacy and Bartlett's Test of Sphericity. Subsequently, CFA was employed to validate the factor structure revealed by EFA. The model's goodness-of-fit was evaluated using several indices: chi-square (χ^2), degrees of freedom (df), χ^2/df ratio, Root Mean Square Error of Approximation (RMSEA), GFI, RFI, CFI, IFI, NFI, and TLI. The internal consistency of the scale and its subdimensions was examined using Cronbach's alpha coefficient. In addition, split-half reliability was calculated to provide further evidence of the scale's stability. To investigate potential response bias, Hotelling's T^2 test was utilized. The level of statistical significance was set at $p < 0.05$ for all analyses, and all tests were interpreted within a 95% confidence interval.

Results

Content Validity

Prior to the pilot implementation of the scale, a content analysis of the items and the overall scale was conducted based on feedback from 10 experts. The Item-Level Content Validity Index (I-CVI) and the Scale-Level Content Validity Index (S-CVI) were calculated as 0.99, demonstrating high consistency in expert evaluations. According to the literature, an I-CVI or S-CVI value above 0.80 is generally considered acceptable. In this study, both values were calculated as 0.99, indicating excellent expert agreement on item relevance.

Construct Validity

In the study, the Kaiser–Meyer–Olkin (KMO) value was found to be 0.963, and Bartlett's test of sphericity was statistically significant ($\chi^2 = 12{,}253.632$, df = 351, $p < 0.001$), indicating that the data were suitable for factor analysis (Table 1). Exploratory Factor Analysis (EFA) conducted on responses from 444 participants revealed a three-factor structure that explained 54.29% of the total variance. The first factor accounted for 41.996% of the variance, the second factor explained 6.873%, and the third factor contributed 5.421% to the total variance (Table 1). The decision to retain three factors was based on the Kaiser criterion (eigenvalues greater than 1) and further supported by the scree plot, which indicated a clear point of inflection after the third factor.As shown in the scree plot (Fig. 1), the point of inflection occurred after the third factor, supporting the retention of a three-factor structure. This result is consistent with the eigenvalues greater than 1 criterion and further confirms the construct validity of the scale.

The Pearson correlation coefficients revealed statistically significant positive relationships among the three subdimensions of the scale ($p < 0.001$). Specifically, the correlation between Factor 1 (Basic Knowledge of Generative Artificial Intelligence) and Factor 2 (Positive Attitudes Toward Generative Artificial Intelligence) was r = 0.692, between Factor 1 and Factor 3 (Concerns About the Impacts of Generative Artificial Intelligence) was r = 0.670, and between Factor 2 and Factor 3 was r = 0.632. These moderate to strong interfactor correlations indicate that while the subdimensions are conceptually distinct, they are meaningfully related, supporting the multidimensional yet cohesive structure of the scale.

During the EFA, items with cross-loadings were reviewed. Only items with a primary loading ≥ 0.40 and a minimum difference of 0.20 between the highest and next loading were retained. This ensured that all items in the final model loaded clearly on a single factor with no significant cross-loading. The factor loadings for the items ranged from 0.458 to 0.795 in the first factor (Basic Knowledge of Generative Artificial Intelligence), 0.478 to 0.805 in the second factor (Positive Attitudes Toward Generative Artificial Intelligence), and 0.457 to 0.685 in the third factor (Concerns About the Impacts of Generative Artificial Intelligence), indicating moderate to strong

Table 1 Results of the exploratory factor analysis ($n = 444$)

		Factor 1	Factor 2	Factor 3
Item 1	I can explain what generative artificial intelligence is in my own words	0.668		
Item 2	I am familiar with generative AI tools (such as ChatGPT, Gemini, DeepSeek, Grok, etc.)	0.786		
Item 3	I know what kinds of tasks generative AI tools can perform	0.795		
Item 4	I have an understanding of how generative AI tools can be used in different areas (e.g., education, content creation, daily life)	0.791		
Item 5	I have used at least one generative AI application at least once	0.681		
Item 6	I use generative AI tools for various purposes (e.g., gathering information, generating text, summarizing)	0.618		
Item 7	I know the steps to follow when using generative AI tools	0.674		
Item 8	I am aware of ethical issues such as copyright and plagiarism when using generative AI tools	0.458		
Item 9	I know that generative AI algorithms can sometimes produce biased or inaccurate information	0.571		
Item 10	I know that information I share when using generative AI tools should be kept confidential	0.556		
Item 11	I am aware that the misuse of generative AI (e.g., creating fake news) can lead to serious societal problems	0.534		
Item 12	I feel confident in learning and using generative AI tools	0.513		
Item 13	I enjoy learning about generative artificial intelligence technologies		0.588	
Item 14	The opportunities provided by generative AI tools excite me		0.678	
Item 15	I have a generally positive attitude toward generative artificial intelligence		0.478	
Item 16	I believe generative artificial intelligence will improve our lives overall		0.525	
Item 17	I would like to use generative AI tools in my classes to better understand the topics		0.552	
Item 18	I am eager to learn more about generative AI on my own		0.805	
Item 19	I would actively participate in a project or club related to generative AI if it were available at school		0.762	
Item 20	I believe gaining generative AI skills can enhance my academic success or creativity		0.604	
Item 21	I would like to see more space dedicated to generative AI education in the school curriculum		0.629	
Item 22	I am concerned that generative artificial intelligence may lead to the disappearance of certain professions in the future			0.685
Item 23	I believe that developments in the field of generative artificial intelligence could impact the career path I choose in the future			0.589
Item 24	I think generative AI technologies are driving significant changes in society			0.676
Item 25	Generative artificial intelligence may alter the ways people communicate and interact with one another			0.672
Item 26	Generative artificial intelligence could negatively affect society in areas such as education and healthcare			0.457
Item 27	I am concerned that generative artificial intelligence may lead to the disappearance of certain professions in the future			0.657
Eigenvalues		11.339	1.856	1.464
Explained variance (%)		41.99%	6.873%	5.421%

KMO = 0.963, Bartlett's Test of Sphericity = 12,253.632, df = 351, $p < 0.001$, Total extraction variance = 54.29%

KMO: Kaiser−Meyer Olkin coefficient

item-factor associations. These findings support the multidimensional construct of the scale and provide robust evidence of its construct validity in measuring perceptions of generative artificial intelligence (see Table 1).

Confirmatory Factor Analysis

CFA was conducted to verify the three-factor structure previously identified through EFA. The results demonstrated

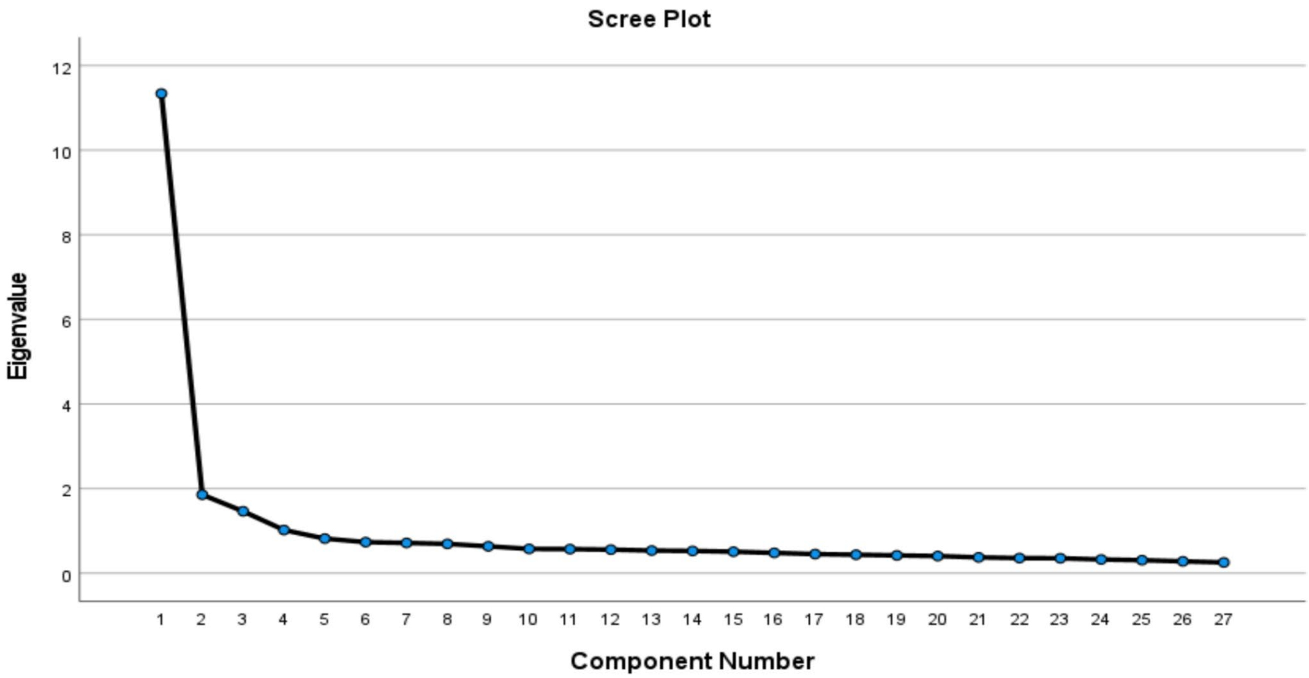

Fig. 1 Scree Plot Displaying the Eigenvalues of the Extracted Components

an acceptable to good fit of the model to the data (Fig. 2). The chi-square value (χ^2) was 1194.547 with 315 degrees of freedom, resulting in a χ^2/df ratio of 3.792, which is within acceptable limits for model fit. The RMSEA was calculated as 0.056, with a 90% confidence interval of [0.053, 0.060], indicating a good model fit. According to conventional criteria, RMSEA values below 0.06 suggest a good fit, and values below 0.08 are considered acceptable. Additional model fit indices further supported the adequacy of the three-factor model. The GFI was 0.905, the RFI was 0.893, and the IFI was 0.927. Moreover, the CFI and the TLI were 0.927 and 0.919, respectively, while the NFI was 0.904. All of these values exceed the commonly accepted threshold of 0.90, indicating that the three-factor model demonstrated strong construct validity and a satisfactory fit to the observed data (see Table 2).

Reliability Analysis

The reliability of the scale was assessed using Cronbach's alpha coefficient and corrected item-total score correlation analyses. The overall Cronbach's alpha for the entire scale was calculated as 0.946, indicating excellent internal consistency. Subscale reliability values were also high, with Cronbach's alpha coefficients of 0.925 for Basic Knowledge of Generative Artificial Intelligence, 0.885 for Positive Attitudes Toward Generative Artificial Intelligence, and 0.822 for Concerns About the Impacts of Generative Artificial Intelligence.

Corrected item-total score correlation coefficients ranged from 0.512 to 0.733, demonstrating moderate to strong relationships between each item and the total score of its corresponding subscale. All values exceeded the commonly accepted threshold of 0.30, supporting the internal consistency and reliability of each item within its designated factor (see Table 3).

To further assess the internal consistency of the instrument, a split-half reliability analysis was conducted for the entire 27-item scale. The Cronbach's alpha coefficient was 0.920 for the first half and 0.888 for the second half. The Spearman–Brown coefficient was calculated as 0.891, while the Guttman split-half coefficient was 0.887. The correlation between the two halves was 0.804, indicating a high level of internal consistency for the overall scale (see Table 4).

The overall Cronbach's alpha value for the entire scale was 0.946, demonstrating excellent internal consistency. Subscale reliability coefficients were also high: 0.925 for the first factor (Basic Knowledge of Generative Artificial Intelligence), 0.885 for the second factor (Positive Attitudes Toward Generative Artificial Intelligence), and 0.822 for the third factor (Concerns About the Impacts of Generative Artificial Intelligence). These findings provide robust evidence for the reliability of the scale and its subdimensions in measuring perceptions related to generative artificial intelligence (see Table 4).

Hotelling's T^2 test was employed to examine the presence of potential response bias within the scale. The analysis yielded a Hotelling's T^2 value of 944.177, corresponding to

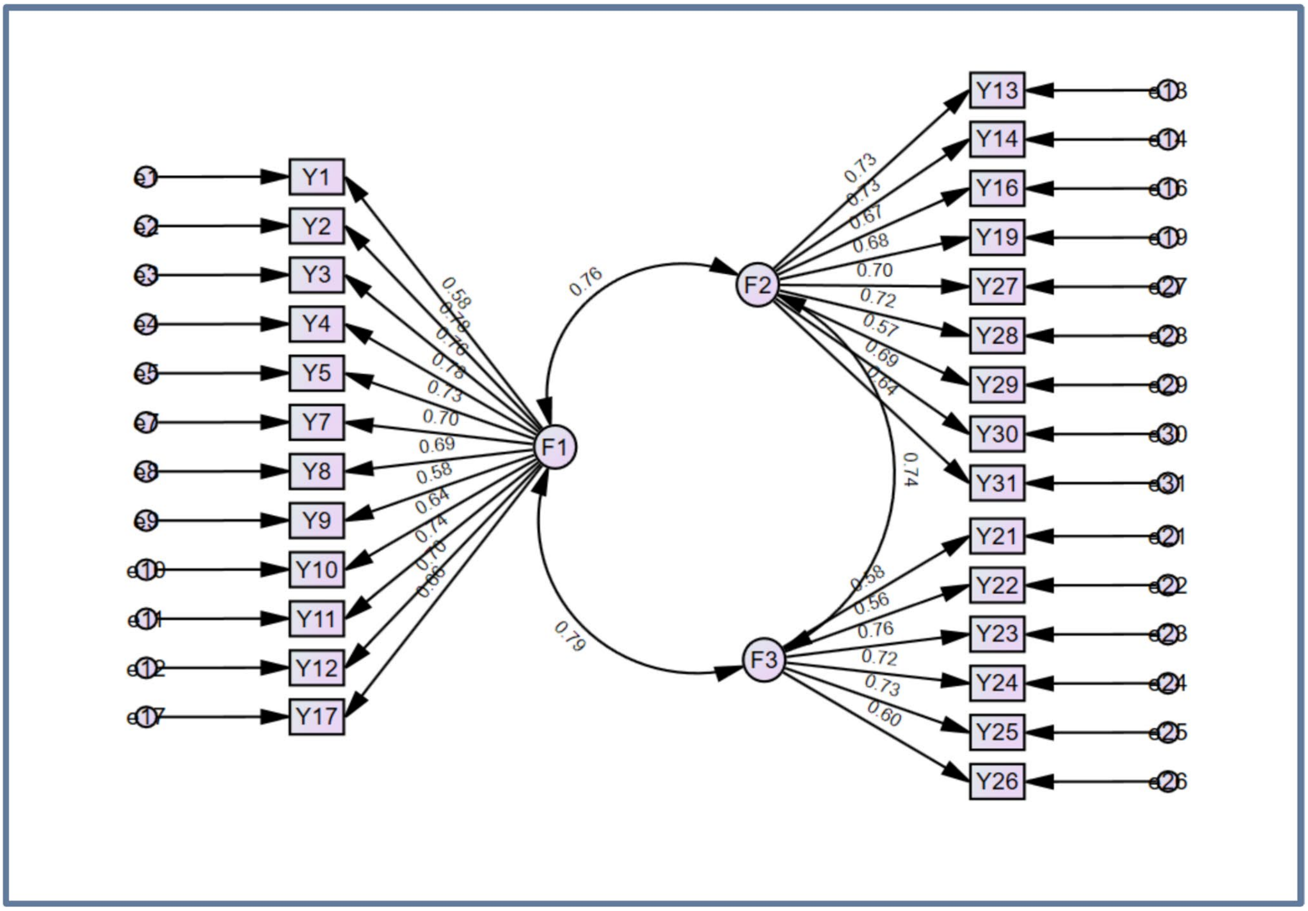

Fig. 2 Path diagram and factor loads of the three-factor models of the scale

Table 2 Model fit indices of the scale ($n = 444$)

	X^2	DF[a]	X^2/DF	RMSEA[b]	GFI[c]	RFI[d]	IFI[e]	CFI[f]	NFI[g]	TLI[h]
Three factor model	1194.547	315	3.792	0.056	0.905	0.893	0.927	0.927	0.904	0.919

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): Trucker–Lewis Index

an F-value of 35.291 and a statistically significant p-value ($p < 0.001$). These findings indicate substantial multivariate differences among item responses, thereby supporting the scale's capacity to discriminate effectively between participants and confirming the absence of uniform response patterns.

Discussion

The application areas of GAI are rapidly expanding, offering significant opportunities, particularly in the field of education. To benefit effectively from this technology, students' level of knowledge, attitudes, and ethical awareness regarding generative AI is of critical importance. It is especially necessary to assess the awareness levels of secondary school students, who are introduced to digital tools at an early age, toward this technology. Accordingly, the present study aims to develop a valid and reliable measurement tool to assess generative AI awareness among secondary school students.

Content validity refers to the extent to which the items in a measurement tool accurately represent the targeted conceptual construct [35]. In this regard, the appropriateness of the item pool created by the researchers was reviewed by a panel of experts. The literature suggests that a minimum agreement rate of 80% at the item level (I-CVI) and the scale level (S-CVI) is considered acceptable for ensuring content

Table 3 Results of the reliability analyses of the scale and correlations of the item total score (444)

Items	Cronbach α	Mean ± SD	Corrected item-total score correlation
1	**Factor 1 = 0.925**	3.38 ± 0.97	0.562
2		4.08 ± 1.02	0.716
3		3.90 ± 0.97	0.706
4		4.01 ± 0.96	0.718
5		4.24 ± 1.06	0.695
6		4.01 ± 1.03	0.688
7		3.69 ± 0.99	0.676
8		3.60 ± 1.06	0.587
9		3.84 ± 1.03	0.625
10		3.95 ± 1.04	0.733
11		4.04 ± 1.05	0.676
12		3.77 ± 0.98	0.682
13	**Factor 2 = 0.885**	3.70 ± 1.07	0.684
14		3.63 ± 1.06	0.655
15		3.68 ± 0.98	0.659
16		3.62 ± 1.04	0.691
17		3.83 ± 1.07	0.654
18		3.50 ± 1.04	0.668
19		3.29 ± 1.08	0.630
20		3.60 ± 1.02	0.512
21		3.55 ± 1.08	0.653
22	**Factor 3 = 0.822**	3.71 ± 1.06	0.534
23		3.55 ± 1.07	0.527
24		3.82 ± 1.03	0.671
25		3.77 ± 1.01	0.634
26		3.92 ± 1.01	0.698
27		3.69 ± 1.04	0.559

Total Cronbach Alpha (α) Value = 0.946

validity [39, 40]. In this study, the final version of the scale, revised based on expert evaluations, consisted of 27 items, and both the I-CVI and S-CVI values were calculated as 0.99. These high values indicate that the scale items clearly and accurately represent the intended conceptual construct and confirm the content validity of the scale.

The second stage of validity analysis involves evaluating construct validity. Construct validity is the process of assessing the extent to which a measurement tool accurately represents the theoretical construct it aims to measure [41, 42]. In scale development research, this is commonly assessed through EFA and CFA. Prior to conducting these analyses, it is crucial to determine the adequacy of the dataset for factor analysis. The KMO coefficient serves this purpose, indicating sampling adequacy. According to established guidelines, KMO values in the range of 0.90–1.00 reflect excellent suitability; 0.80–0.89 very good; 0.70–0.79 acceptable; 0.60–0.69 marginal; and 0.50–0.59 poor [43]. In this study, the KMO value was found to be 0.963, indicating that the sample was highly suitable for analysis. Additionally, the significant *p*-value ($p < 0.005$) obtained from Bartlett's test of sphericity confirmed that the variables were sufficiently correlated and that the dataset was appropriate for factor analysis [38]. All these findings support the suitability of the dataset for factor analysis within the scope of this research.

EFA is used to determine the factor structure of a scale [44]. According to the scale development literature, explaining 40% to 60% of the total variance is generally considered an acceptable level of explanatory power [45]. In the present study, the factor loadings of the items in the three subdimensions ranged from 0.458 to 0.795, and the total variance explained was within the acceptable limits indicated in the literature.

Following the EFA, CFA was performed to verify the factor structure of the scale, and model fit indices were examined [46]. Various threshold values for CFA fit indices are recommended in the literature. For example, Konan et al. (2015) stated that a χ^2/df ratio below 5, and CFI, NNFI, and AGFI values above 0.90 indicate strong model fit, while CFI ≥ 0.85 and AGFI ≥ 0.80 reflect acceptable fit [47]. Wong et al. (2012) considered χ^2/df ratios between 2 and 5 as indicators of good fit and defined an RMSEA value close to 0.06 as good and values near 0.00 as excellent [48]. For GFI and CFI, values ≥ 0.90 are minimally acceptable, ≥ 0.95 indicate good fit, and values approaching 1.00 reflect ideal fit. In this study, the obtained fit indices aligned with these criteria, indicating that the model provided an overall adequate fit.

Based on the obtained fit indices, the three-factor structure of the scale was validated, and the model demonstrated construct validity. The reliability of a measurement tool refers to its ability to produce consistent results under the same conditions [49, 50]. Two key indicators are commonly used in reliability assessments: consistency over time (stability) and internal consistency of items. To test stability, it is recommended to re-administer the instrument after a certain period [51]. The literature suggests that a correlation

Table 4 Results of reliability analysis of scale and subdimensions ($n = 888$)

Subdimensions	Cronbach α	Split into two halves				
		First half Cronbach α	Second half Cronbach α	Spearman-Brown	Guttman split-half	Correlation between two halves
Total scale	0.946	0.920	0.888	0.891	0.887	0.804

coefficient of at least 0.70 is considered acceptable for test–retest reliability [52, 53].

To assess the stability of the scale over time, it was re-administered to the same sample following a two-week interval. Pearson correlation analysis was employed to examine the relationship between the two sets of scores. The results demonstrated that the scale produces stable and consistent outcomes across time points, thereby confirming its temporal reliability and its ability to reflect the underlying construct. In addition, internal consistency was evaluated by analysing the correlations between each item and the total score. The analysis yielded a high Cronbach's alpha coefficient for the entire scale, indicating strong internal reliability and supporting the scale's robustness as a psychometric instrument.

In addition to demonstrating the scale's strong psychometric properties, this study makes several important theoretical and practical contributions to the field. First, the GAIAS responds to a pressing gap in the literature by offering a comprehensive tool specifically designed for secondary school students, a demographic increasingly exposed to generative AI tools but underrepresented in empirical research. Unlike general AI literacy scales that tend to focus on functional knowledge or are tailored to university populations [34], GAIAS captures the multidimensional nature of generative AI awareness, including students' basic knowledge, affective orientations, and ethical concerns.

The development of this scale is particularly timely, given the growing emphasis on students' preparedness for AI-integrated learning environments. As emphasized by Ng et al. (2024), fostering ethical, cognitive, and behavioural readiness in adolescents is essential for promoting responsible engagement with generative technologies [29]. The three-factor structure of GAIAS aligns with contemporary frameworks of AI literacy and digital agency [14, 54], and may serve as a foundation for future longitudinal studies exploring how awareness evolves across developmental stages.

Moreover, the scale has direct implications for educational practice. Educators and policymakers can use GAIAS to assess students' readiness, identify potential misconceptions, and design targeted interventions that strengthen digital literacy and ethical reasoning in the context of generative AI. For instance, students who demonstrate high levels of technical familiarity but low ethical awareness could benefit from curriculum components emphasizing responsible use, academic integrity, and data privacy [11, 21].

Finally, this study contributes to the broader discourse on the psychological and social implications of AI. In light of recent findings that link increased AI usage to concerns such as reduced self-regulation and heightened anxiety [55, 56], the scale developed here enables the early identification of students at risk of cognitive overload or ethical disengagement. As such, GAIAS offers a valuable measurement framework for scholars and practitioners seeking to understand and shape the next generation's relationship with generative technologies.

In conclusion, understanding the impact of generative artificial intelligence technologies in education requires a reliable assessment of students'knowledge, attitudes, and awareness regarding these technologies. The scale developed in this study not only possesses valid and reliable psychometric properties but also enables early identification of students'psychosocial perceptions and potential cognitive-technological challenges related to generative AI. As emphasized in the literature, it is critically important to recognize issues such as anxiety, stress, and uncertainty arising from technological transformations, particularly among children and adolescents, in a timely manner to ensure academic success and social adaptation [55, 56]. In this regard, the Generative Artificial Intelligence Awareness Scale provides a scientifically grounded means of assessing the potential psychosocial impacts of students' interactions with the digital world and contributes meaningfully to the academic literature.

Limitations and future studies

Although this study aimed to develop a valid and reliable scale to assess secondary school students' awareness of GAI, it has several noteworthy limitations that also present opportunities for further research.

First, the sample was limited to students aged 14–18 from four secondary education institutions in Türkiye. This poses constraints on the generalizability of the findings to other educational systems, technological infrastructures, and sociocultural contexts. While the current study mentions the importance of cross-cultural validation, future studies should go further by identifying which specific cultural dimensions such as public discourse around AI ethics, national digital education policies, or collective vs. individualistic learning styles may affect students' awareness and engagement with GAI tools. For instance, students from countries with high levels of AI integration in education may exhibit different patterns of awareness than those in low-resource or low-exposure settings.

Second, although the scale was developed in Turkish and showed strong psychometric properties, its applicability to broader populations requires linguistic and cultural adaptation. Future studies should undertake rigorous translation and back-translation processes, followed by measurement invariance testing in other languages such as English, Spanish, or Mandarin. This will ensure the semantic, conceptual, and functional equivalence of the scale across diverse populations.

Third, the present study focused primarily on cognitive and attitudinal components of awareness (i.e., knowledge, positive attitude, concern). However, affective and social dimensions such as emotional regulation, peer influence, and collaborative decision-making—were not explicitly incorporated. The literature increasingly acknowledges that awareness of emerging technologies like GAI is not formed in isolation but is influenced by socio-emotional interactions and contextual learning environments [53, 54]. Therefore, future research could extend this work by integrating perspectives from social cognitive theory and sociocultural theory, exploring how social modeling, peer norms, and cultural narratives shape adolescents' ethical reasoning and technological engagement.

Fourth, although the need for "more comprehensive models" was acknowledged, the current recommendations for future studies remain abstract. We recommend that future researchers employ mixed-methods designs to capture both quantitative and qualitative nuances of GAI awareness; longitudinal studies to examine how awareness evolves over time with increased exposure; and multidimensional scaling or network analysis to explore the interplay between cognitive, emotional, and behavioral variables in GAI use. Furthermore, new subdimensions such as privacy concerns, fear of surveillance, or AI dependency could be tested in relation to students' psychological well-being and academic performance.

Lastly, while Hotelling's T^2 test was used to examine potential response bias, it should be noted that the test assumes multivariate normality, which may not strictly hold with Likert-type data. This limitation is acknowledged, and the primary purpose of the analysis was to identify systematic response patterns.

By addressing these limitations and extending the conceptual scope, future studies can refine the measurement of GAI awareness and contribute to the development of ethically grounded, culturally responsive, and developmentally appropriate educational strategies in the age of generative AI.

Conclusion

This study aimed to develop a psychometrically sound and developmentally appropriate instrument to assess GAI awareness among secondary school students. Given the limited number of validated tools that address GAI awareness particularly within adolescent populations this research fills a critical gap in the literature. Unlike general AI literacy or digital literacy scales, the GAIAS specifically captures dimensions unique to generative AI, such as creative engagement, ethical concerns, and usage behavior, thereby contributing to the theoretical and empirical differentiation of GAI-related constructs.

The academic significance of this study lies not only in the development of a new measurement tool but also in its potential to support future interdisciplinary research on adolescents' cognitive, ethical, and behavioral responses to emerging AI technologies. The scale offers a framework for understanding how students engage with GAI in ways that are distinct from traditional AI systems, paving the way for more targeted investigations into digital ethics, responsible AI use, and human–AI interaction.

From a policy and educational standpoint, the scale provides actionable insights for curriculum designers, educators, and decision-makers. For example, the GAIAS can be integrated into digital literacy education modules to evaluate students'readiness for AI enhanced learning environments. It can also inform the development of classroom-based interventions that promote ethical reasoning and critical thinking in AI usage, particularly in subjects such as computer science, citizenship education, and media literacy.

In sum, the GAIAS offers a timely and contextually relevant tool for both researchers and practitioners. As GAI tools continue to shape the digital landscape of education, assessing and nurturing awareness among younger learners becomes essential for fostering not just technological competence but also ethical responsibility and informed participation in AI-driven societies.

Authors' contributions Conceptualization: R.S.Ş., Ö.Ö., S.Ç.Ö. and S.D.S. Literature search and screening: R.S.Ş., Ö.Ö., S.Ç.Ö. and S.D.S.. Methodology: R.S.Ş., Writing the original draft, reviewing, and editing: R.S.Ş., Ö.Ö., S.Ç.Ö. and S.D.S. authors read and approved the final version of the manuscript.

Funding This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

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

Declarations

Ethical approval Ethical approval for the study was obtained from the Scientific Research and Publication Ethics Committee of Artvin Çoruh University (Approval No: E-18457941–050.99–171843). In addition, permission to conduct the study was obtained from the Provincial Directorate of National Education (Approval No: MEB. TT.2025.022666.01).

Consent for participate Not applicable.

Competing interests The authors declare no competing interests.

References

1. Mao J, Chen B, Liu JC (2024) Generative artificial intelligence in education and its implications for assessment. TechTrends 68(1):58–66. https://doi.org/10.1007/s11528-023-00855-9
2. Traylor DO, Kern KV, Anderson EE, Henderson R (2025) Beyond the Screen: The impact of generative artificial intelligence (AI) on patient learning and the patient-physician relationship. Cureus 17(1):52913. https://doi.org/10.7759/cureus.52913

3. Ooi KB, Tan GWH, Al-Emran M, Al-Sharafi MA, Capatina A, Chakraborty A et al (2025) The potential of generative artificial intelligence across disciplines: Perspectives and future directions. J Comput Inf Syst 65(1):76–107. https://doi.org/10.1080/08874417.2023.2225096
4. Wu D, Zhang J (2025) Generative artificial intelligence in secondary education: applications and effects on students' innovation skills and digital literacy. PLoS ONE 20(5):e0323349. https://doi.org/10.1371/journal.pone.0323349
5. Creely E, Blannin J (2025) Creative partnerships with generative AI: possibilities for education and beyond. Think Ski Creat 56:101727. https://doi.org/10.1016/j.tsc.2024.101727
6. Habib S, Vogel T, Anli X, Thorne E (2024) How does generative artificial intelligence impact student creativity? J Creat 34(1):100072. https://doi.org/10.1016/j.crea.2024.100072
7. Maity S, Deroy A. (2024) Generative AI and its impact on personalized intelligent tutoring systems. arXiv Prepr. arXiv:2410.10650. https://doi.org/10.48550/arXiv.2410.10650
8. Dhagare MRP (2024) Generative AI and education: A symbiotic relationship. Int J Res Appl Sci Eng Technol. 12(11):1042–1045. https://doi.org/10.22214/ijraset.2024.61884
9. Wirzal MDH, Nordin NAHM, Abd NS, Halim M. (2024) Generative AI in science education: A learning revolution or a threat to academic integrity? A bibliometric analysis. J Educ Res Stud e-Saintika. 8(3):319–351. https://doi.org/10.36312/jes.v8i3.1456
10. Francis NJ, Jones S, Smith DP (2025) Generative AI in higher education: Balancing innovation and integrity. Br J Biomed Sci 81:14048. https://doi.org/10.1080/09674845.2024.2334568
11. Hill C, Hargis J (2024) An ethics module on academic integrity and generative AI. New Dir Teach Learn. https://doi.org/10.1002/tl.20612
12. Creely E (2024) Exploring the role of generative AI in enhancing language learning: opportunities and challenges. Int J Chang Educ 1(3):158–167. https://doi.org/10.47852/bonviewIJCE42022495
13. Tadimalla SY, Maher M Lou. (2024) AI literacy for all: Adjustable interdisciplinary socio-technical curriculum. arXiv Prepr. arXiv:2409.10552. https://doi.org/10.48550/arXiv.2409.10552
14. Biagini G. (2024) Assessing the assessments: Toward a multidimensional approach to AI literacy. Media Educ. 15(1):91–101. https://doi.org/10.36253/mediaed-15077
15. Ali O, Murray PA, Momin M, Dwivedi YK, Malik T (2024) The effects of artificial intelligence applications in educational settings: challenges and strategies. Technol Forecast Soc Change 199:123076. https://doi.org/10.1016/j.techfore.2023.123076
16. Kanbach DK, Heiduk L, Blueher G, Schreiter M, Lahmann A (2024) The GenAI is out of the bottle: generative artificial intelligence from a business model innovation perspective. Rev Manag Sci 18(4):1189–1220. https://doi.org/10.1007/s11846-023-00696-z
17. Zhao C, Du H, Niyato D, Kang J, Xiong Z, Kim DI, Shen X, Letaief KB (2025) Enhancing physical layer communication security through generative AI with mixture of experts. IEEE Wirel Commun 32(3):176–184. https://doi.org/10.1109/MWC.001.2400150
18. Lang Q, Wang M, Yin M, Liang S, Song W (2025) Transforming education with generative AI (GAI): key insights and future prospects. IEEE Trans Learn Technol 18:230–242. https://doi.org/10.1109/TLT.2025.3537618
19. Al Naqbi H, Bahroun Z, Ahmed V (2024) Enhancing work productivity through generative artificial intelligence: a comprehensive literature review. Sustainability 16(3):1166. https://doi.org/10.3390/su16031166
20. Naranjo JE, Llumiquinga MM, Vaca WD, Espin CX (2025) Generative AI vs. traditional databases: insights from industrial engineering applications. Publications 13(2):14. https://doi.org/10.3390/publications13020014
21. Kwon J (2024) Enhancing ethical awareness through generative AI literacy: a study on user engagement and competence. Edelweiss Appl Sci Technol 8(6):4136–4145. https://doi.org/10.55214/25768484.v8i6.2904
22. Lo CK (2023) What is the impact of ChatGPT on education? A rapid review of the literature. Educ Sci 13(4):410. https://doi.org/10.3390/educsci13040410
23. Mohammad B, Supti T, Alzubaidi M, Shah H, Alam T, Shah Z, et al. (2023) The pros and cons of using ChatGPT in medical education: A scoping review. In: Healthc Transform with Informatics Artif Intell. 644–647. https://doi.org/10.3233/SHTI230580
24. Dungo CAB, Beltran ZLE, Declaro BC, Dela-Cruz JJC, Viray RU. (2025) Students' level of awareness on the environmental implications of generative AI. J Educ Sci Environ Heal. 11(2):93–107. https://doi.org/10.12973/JESEH.2025.11.2.6
25. Zhang P, Tur G (2024) A systematic review of ChatGPT use in K-12 education. Eur J Educ 59(2):e12599. https://doi.org/10.1111/ejed.12599
26. Mogavi RH, Deng C, Kim JJ, Zhou P, Kwon YD, Metwally AHS et al (2024) ChatGPT in education: a blessing or a curse? A qualitative study exploring early adopters' utilization and perceptions. Computers in Human Behavior: Artificial Humans 2(1):100027. https://doi.org/10.1016/j.chbah.2023.100027
27. Wen Y, Zhao X, Li X, Zang Y (2025) Attitude mining toward generative artificial intelligence in education: The challenges and responses for sustainable development in education. Sustainability 17(3):1127. https://doi.org/10.3390/su17031127
28. Garofalo SG, Farenga SJ (2025) Science teacher perceptions of the state of knowledge and education at the advent of generative artificial intelligence popularity. Sci Educ 34(2):893–912. https://doi.org/10.1002/sce.21838
29. Ng DTK, Wu W, Leung JKL, Chiu TKF, Chu SKW (2024) Design and validation of the AI literacy questionnaire: The affective, behavioural, cognitive and ethical approach. Br J Educ Technol 55(3):1082–1104. https://doi.org/10.1111/bjet.13386
30. Chukhno O. (2024) Teachers and learners' perspectives on the use of generative AI in foreign language education. Mod Inf Technol Innov Methodol Educ Prof Train Methodol Theory Exp Probl. (73):53–60. https://doi.org/10.31652/2412-1142-2024-73-53-60
31. Wood D, Moss SH (2024) Evaluating the impact of students' generative AI use in educational contexts. J Res Innov Teach Learn 17(2):152–167. https://doi.org/10.1108/JRIT-12-2023-0096
32. Zhang H, Lee I, Ali S, DiPaola D, Cheng Y, Breazeal C (2023) Integrating ethics and career futures with technical learning to promote AI literacy for middle school students: An exploratory study. Int J Artif Intell Educ 33(2):290–324. https://doi.org/10.1007/s40593-022-00323-9
33. Zhang H, Perry A, Lee I. (2024) Developing and validating the artificial intelligence literacy concept inventory: An instrument to assess artificial intelligence literacy among middle school students. Int J Artif Intell Educ. 1–41. https://doi.org/10.1007/s40593-024-00360-w
34. Zhang H, Perry A, Lee I (2025) Developing and validating the artificial intelligence literacy concept inventory: an instrument to assess artificial intelligence literacy among middle school students. Int J Artif Intell Educ 35(1):398–438. https://doi.org/10.1007/s40593-024-00398-x
35. DeVellis RF, Thorpe CT. (2021) Scale development: Theory and applications. Sage Publications.
36. Karagöz Y (2018) SPSS and AMOS 23 applied statistical analysis. Nobel Publishing, Ankara
37. Polit DF, Beck CT (2006) The content validity index: are you sure you know what's being reported? Critique and recommendations. Res Nurs Health 29(5):489–497. https://doi.org/10.1002/nur.20147

38. Şencan H (2005) Reliability and validity in social and behavioral measurements. Seçkin Publishing, Ankara
39. Çapık C, Gözüm S, Aksayan S. (2018) Kültürlerarası ölçek uyarlama aşamaları, dil ve kültür uyarlaması: Güncellenmiş rehber. Florence Nightingale J Nurs. 26(3):199–210. https://doi.org/10.26650/FNJN18061
40. Yusoff MSB. (2019) ABC of content validation and content validity index calculation. Educ Med J. 11(2):49–54. https://doi.org/10.21315/emj2019.11.2.6
41. Seçer İ. (2018) Psikolojik test geliştirme ve uyarlama süreci: SPSS ve LISREL uygulamaları. Ankara: Anı Yayıncılık.
42. Tsang S, Royse CF, Terkawi AS (2017) Guidelines for developing, translating, and validating a questionnaire in perioperative and pain medicine. Saudi J Anaesth 11(Suppl 1):S80–S89. https://doi.org/10.4103/sja.SJA_203_17
43. Yaşlıoğlu MM. (2017) Factor analysis and validity in social sciences: Application of exploratory and confirmatory factor analyses. Istanbul Univ J Sch Bus. 46(Special issue):74–85.
44. Karagöz Y. (2021) SPSS ve AMOS uygulamalı nicel-nitel-karma bilimsel araştırma yöntemleri ve yayın etiği. Nobel Akademik Yayıncılık.
45. Deniz Doğan S, Arslan S (2019) Validity and reliability of Turkish version of the burn-specific pain anxiety scale. J Burn Care Res 40(6):818–822. https://doi.org/10.1093/jbcr/irz084
46. Brown TA. (2015) Confirmatory factor analysis for applied research. Guilford Publications.
47. Konan N, Demir H, Karakuş M (2015) A study of Turkish adaptation of executive servant leadership scale into Turkish. Electron Int J Educ Arts Sci 1(1):135–155
48. Wong AWK, O'Sullivan D, Strauser DR (2012) Confirmatory factor analytical study of the revised developmental work personality scale. Meas Eval Couns Dev 45(4):270–291. https://doi.org/10.1177/0748175612451741
49. Beesdo-Baum K, Knappe S (2012) Developmental epidemiology of anxiety disorders. Child Adolesc Psychiatr Clin 21(3):457–478. https://doi.org/10.1016/j.chc.2012.05.001
50. Bonett DG, Wright TA (2015) Cronbach's alpha reliability: interval estimation, hypothesis testing, and sample size planning. J Organ Behav 36(1):3–15. https://doi.org/10.1002/job.1960
51. Morgado FFR, Meireles JFF, Neves CM, Amaral ACS, Ferreira MEC (2017) Scale development: ten main limitations and recommendations to improve future research practices. Psicol Reflex Crit 30(1):3. https://doi.org/10.1186/s41155-016-0057-1
52. Bernstein IH. (2010) Psychometric theory. Tata McGraw-Hill Education.
53. Chakrabartty SN (2013) Best split-half and maximum reliability. IOSR J Res Method Educ 3(1):1–8. https://doi.org/10.9790/7388-0310108
54. Code J (2020) Agency for learning: intention, motivation, self-efficacy and self-regulation. Front Educ. https://doi.org/10.3389/feduc.2020.00019
55. Grist R, Croker A, Denne M, Stallard P (2019) Technology-delivered interventions for depression and anxiety in children and adolescents: A systematic review and meta-analysis. Clin Child Fam Psychol Rev 22:147–171. https://doi.org/10.1007/s10567-018-0276
56. Meates J. (2020) Problematic digital technology use of children and adolescents: Psychological impact. Teachers and Curriculum. 20(1):51–62. https://doi.org/10.15663/tandc.v20i1.354