QUBAHAN ACADEMIC JOURNAL
VOL. 5, NO. 2, April 2025
https://doi.org/10.48161/qaj.v5n2a1672

Factorial Structure and Measurement Invariance of the General Atitudes Toward Artificial Intelligence Scale for University Students

Mohammed Mousa Al-Shumrani [1*]

[1] Department of Psychology, College of Education, Taif University, Al-Hawiyah, Taif 21944, Saudi Arabia.

* **Corresponding author:** m.m.m.a1313@hotmail.com.

ABSTRACT: The study aims to investigate the factorial structure (FS) and measurement invariance (MI) of the General Attitudes Toward Artificial Intelligence Scale (GAAIS) on a sample of students from Taif University. The instrument used was the GAAIS developed by [1] and translated by the researcher. The scale was administered to a sample of 461 university students. The study results indicated that the scale's structure aligned with the confirmatory factor analysis model, with fit indices demonstrating good fit: ($\chi^2/df=611.77/169=3.62$), Good of fit Index (GFI) = 0.855, Tucker-Lewis index (TLI) = 0.837, and Root Mean square of Approximation (RMSEA) = 0.075. These results confirm the construct validity and the scale's suitability for the data. The composite reliability coefficients were 0.828 and 0.869, and Cronbach's alpha reliability coefficients were 0.832 and 0.869 for the positive sub-dimensions (opportunities, benefits, societal and personal utility, and positive emotions) and the negative sub-dimensions (fears and negative emotions), respectively. Additionally, the study confirms configural, metric, scalar, and residual invariance across the male and female study groups. The results confirmed that the scale has high reliability and validity and is unbiased regarding the gender variable. The study also confirms the suitability of the instrument for measuring general attitudes toward artificial intelligence (AI) among university students.

Keywords: attitudes, factorial structure, measurement invariance, artificial intelligence.

I. INTRODUCTION

AI is considered one of the most significant innovations that has profoundly impacted all aspects of human life. AI technologies are used in various fields such as medicine, education, and industry, and have become a fundamental cornerstone for developing innovative solutions and enhancing efficiency across numerous sectors [2]. AI technology is a combination of intelligent software and a branch of computer technology that focuses on interactive smart software systems by performing multiple functions and tasks that simulate human work. It scientifically contributes to many fields, as AI plays an important and growing role in various aspects of life and is considered an essential element in human life [3]. With the rapid advancement of AI technologies and their widespread adoption across various fields of life, understanding societal attitudes toward this modern technology has become essential. Therefore, individuals' attitudes toward using these technologies are influenced by various perspectives on this technology [4].

With the penetration of AI into all aspects of life, university students need to acquire knowledge and skills and identify the necessary positive and negative attitudes to adapt to the challenges posed by AI [5]. On the other hand, Cai et al. [6] indicate that although AI has contributed to driving social changes, it has not been fully utilized by the public. One of the obstacles to this limited use may be the extent to which users accept AI.

Chen et al. [7] emphasize the importance of having reliable scientific measurement tools that help identify various positive and negative attitudes toward this technology. Societies that frequently use modern technology tend to develop positive attitudes toward it, while those with lower usage levels tend to be more reserved. This necessitates measurement tools capable of ensuring MI to confirm that differences reflect real outcomes rather than being influenced by the measurement tool itself.

Research also indicates that factor analysis and MI play a crucial role in ensuring the validity and reliability of the measurement tool. Venkatch and Davis [8] assert that to ensure the generalizability of results, it is essential to verify the validity and reliability of measurement tools and to confirm their MI across different variables such as gender, social status, and economic background. Additionally, Davis [9] emphasizes that understanding individuals' attitudes toward modern technologies can help predict the extent to which they adopt and use these technologies. Concerns have emerged, such as job security, AI surveillance, and the misuse of personal data. As a result, people may have both positive and negative attitudes toward AI depending on its technology. Identifying the factors positively or negatively associated with university students' attitudes toward AI can help leverage AI technology in planning and design, ultimately supporting learning and skill acquisition for university students. Therefore, there is a need to focus on measuring attitudes toward AI technology, making it essential to have measurement tools to assess students' attitudes toward AI and determine the factors influencing these positive and negative attitudes. Moreover, verifying the validity of a scale used in different environments can facilitate the measurement of attitudes by providing a means to assess university students' perspectives on AI [5].

To understand user responses to technology, it is necessary to use high-quality measures when evaluating attitudes toward AI. Scientific efforts have produced several scales for assessing attitudes toward AI, including the GAAIS, which provides two subscales that capture both positive and negative aspects of AI [10]. Le [11] notes that individuals differ in their evaluation of the benefits and risks associated with AI, leading to varying attitudes toward it. Therefore, understanding general attitudes toward AI is crucial for its integration and acceptance, as AI plays a significant and vital role in helping students access information, gather data, solve problems, make decisions, and improve skills. Furthermore, Petric and Atanasova [12] emphasize that measurement tools must be valid and reliable, relying on high-quality information in the assessment process. Protzko [13] highlights the importance of ensuring measurement invariance to confirm that a scale measures the same construct across different groups (e.g., gender). Clawson et al. [14] point out the limited number of studies addressing measurement invariance. Based on this perspective, the current study focuses on examining students' attitudes toward AI and verifying the measurement invariance of the scale across the gender variable.

1. STATEMENT OF THE PROBLEM

Attitudes serve as a key factor in determining individuals' perspectives on AI technology. Students, being one of the essential pillars of the academic educational process. Chai et al. [5] emphasize that AI technology is bringing about a radical transformation in all fields and highlight the need to study university students' perceptions and attitudes toward AI. Since university students are among the most affected by AI technology and its various applications especially after graduation, when they enter a rapidly changing job market influenced by AI it is essential to develop or standardize measurement tools to help determine the appropriate stance toward this technology. Kline [15] suggests that one suitable method for verifying the FS of a scale is factor analysis, which reflects the scale's underlying structure.

Additionally, ensuring MI is crucial, as Vandenberg and Lance [16] indicate that if MI is not achieved, the results may be biased, making it difficult to compare groups based on variables such as gender, economic status, or cultural background. Given the acceptance of the GAAIS, it has been used in multiple studies [17] [1] and has been translated into Turkish, as seen in the study by [17]. These studies have demonstrated its validity and reliability through confirmatory factor analysis and the calculation of Cronbach's alpha for reliability, providing a measure of attitudes toward AI in different environments. Currently, there are no existing scales to measure university students' attitudes toward AI in Saudi Arabia. Therefore, developing an Arabic version of the GAAIS is essential for researchers in the field of human-computer interaction.

However, since no Arabic version of the scale exists, its characteristics were examined in terms of FS and MI across the gender variable. Therefore, the current study seeks to answer the following questions:

- What are the fit indices of the GAAIS about the confirmatory factor analysis model among students at Taif University?
- What are the indications of MI for the GAAIS based on the gender variable?
- The current study aims to verify the construct validity and MI of the GAAIS on a sample of students at Taif University.

2. OBJECTIVES OF THE STUDY

The current study aims to enhance our understanding of students' attitudes toward AI by contributing to the growing body of literature on this topic, with a focus on the context of Saudi universities. This is achieved by adapting the Arabic version of the GAAIS and verifying its construct validity and measurement invariance on a sample of students from Taif University, providing a reliable and valid scale applicable to Saudi university students.

3. SIGNIFICANCE OF THE STUDY

This study's significance stems from its contribution to developing measurement tools that enhance the reliability of future research on attitudes toward AI. Providing an equitable scale enables researchers and decision-makers to gain an accurate understanding of individuals' and communities' perspectives on AI, thereby aiding in the development of awareness strategies or policies that support the positive adoption of this technology. (i) The current study addresses measuring general attitudes toward AI, a relatively new concept in Arab culture. (ii) It also helps researchers and practitioners provide a measurement tool that can be used with different variables in conducting research and studies. (iii) A valuable and practical research tool that enables researchers to conduct cross-cultural comparisons of attitudes toward AI. (iv) In addition, it provides a valuable and practical research tool that enables researchers to conduct cross-cultural comparisons of attitudes toward AI.

II. THEORETICAL FRAMEWORK AND PREVIOUS RESEARCH

1. CONCEPT OF ARTIFICIAL INTELLIGENCE

The concept of AI was first introduced in the 1950s, after which numerous studies were conducted on it. With the development of other technologies and further research into AI technology, experts and researchers in this field developed an expert system for AI by the 1970s. With the advancement of computer systems, AI technology witnessed even greater development [18]. In recent years, AI has gained widespread popularity among both academic circles and the general public. It is often credited with positive effects across various fields. On the other hand, there is a negative perspective and growing concern about its potentially harmful impact on individuals and society, respectively [19].

AI technology includes a set of techniques, skills, methods, and processes that create AI systems or machines that may replace humans or enhance their abilities in performing intelligent tasks. AI is defined as the system's ability to correctly interpret external data, learn from it, and use these lessons to achieve goals and tasks through flexible adaptation [5]. It is also defined as the simulation of human intelligence in machines capable of performing specific tasks, such as problem-solving or learning, mimicking human capabilities [20]. Additionally, it is described as the simulation of human intelligence in machines, encompassing several techniques that use learning, logical thinking, and performing tasks that typically require human perception [11].

The term artificial intelligence is used to describe technologies capable of performing tasks that may require human intelligence, the ability to learn and adapt, and the capacity to handle massive amounts of data [10]. AI technology is considered a branch of computer science and refers to computer-based simulation technology that mimics human thinking and behavior patterns, enabling computers to possess response capabilities similar to human responses [20]. With the opportunities and challenges arising from AI developments and its integration into society, understanding AI has become a major concern [21]. It has also become an integral part of many contemporary technologies, such as smart devices, social media platforms, and robots. At the same time, research shows that many people feel concerned about the potential of these technologies [10].

2. GENERAL ATTITUDES TOWARD ARTIFICIAL INTELLIGENCE

AI is a branch of computer science that enables computers to perform tasks typically carried out by humans in an intelligent manner, often with greater speed and efficiency. Attitude refers to an internal mental state that represents individuals' tendencies and responses toward specific subjects, which can be expressed as acceptance or rejection, or as positive or negative reactions.

Lee [23] explains that general attitudes refer to an individual's responses based on past experiences, encompassing beliefs, values, opinions, and impressions influenced by perceived benefits or fears associated with AI technology. Similarly, Sun et al. [24] highlight that while some individuals view this technology positively, seeing it as a tool for enhancing productivity, excellence, and creativity, others perceive it with negativity and fear, concerned about its impact on their future lives. Parasuraman [25] suggests that individuals' attitudes toward technology are influenced by key psychological variables, including optimism, which reflects the belief that diverse technologies provide greater control, efficiency, and flexibility, as well as negative emotions, which are viewed as obstacles to students' attitudes toward AI. Saklaki and Gardikiotis [21] emphasize that the general acceptance of AI usage and its integration into daily life is heavily reliant on individuals' attitudes toward AI, specifically their evaluations of its potential benefits or drawbacks. Research in psychology and social sciences has shown that attitudes are strong indicators of behavior. This makes studying general attitudes toward AI essential for understanding students' perspectives and developing effective policies related to it.

3. MEASURING ATTITUDES TOWARD ARTIFICIAL INTELLIGENCE

Attitudes play a crucial role in the acceptance of technology, which largely depends on the perceived benefits and ease of use. Davis [9] states that measuring attitudes toward computer technology and AI helps assess users' acceptance of these technologies and identify factors that may lead to negative attitudes. Kline [26] also highlights the use of various measurement tools to determine attitudes toward different subjects, including computer technology. Effective measurement tools must undergo several steps, beginning with identifying dimensions and items through a review of previous literature, followed by application to a sample and verification of the scale’s properties. Confirmatory or exploratory factor analysis can be used to assess the FS and reliability of the scale.

4. FACTOR STRUCTURE OF THE SCALE

The FS is determined using factor analysis, which examines the relationships or correlations between scale items and reduces them into a smaller number of components called factors. Byrne [27] states that the GAAIS includes several dimensions: the benefit dimension, which consists of statements measuring the perceived advantages; the concern dimension, which includes statements related to fears and risks; and the ethics dimension, which comprises items addressing the ethical aspects of using this technology.

5. MEASUREMENT INVARIANCE

Measurement Invariance (MI) is a statistical method that reflects the ability of a scale to measure the same concept across different groups based on variables such as gender, economic status, or social status. Brown [28] emphasizes the importance of MI to ensure the appropriateness of a scale for use with different groups. Similarly, Cheung and Rensvold [29] assert that achieving MI is a necessary condition for making valid comparisons between groups regarding a specific trait.

Protzko [13] suggests that MI refers to studying the psychometric properties of a multi-item scale, either across different groups of participants or within the same group over time. The process of testing measurement invariance involves applying the same scale to different groups of individuals (e.g., gender) under the same conditions, and then ensuring that the factor structure remains the same across both groups, so it is identical. Testing measurement invariance is an important step to ensure fairness and accuracy in testing. Often, when measurement invariance is achieved, the scale measures the same construct across all groups. Ding et al. [30] also emphasize the necessity of conducting a measurement invariance test to ensure the quality of measurement tools and to ensure that these tools, under different conditions, provide equivalent representations of the same structure.

6. LEVELS OF MEASUREMENT INVARIANCE

MI is tested across several levels from the simplest (configural invariance) to more complex levels. These levels include: (1) Configural Invariance: This is the simplest level of MI, also referred to as "general invariance." It means that the FS, such as the number of factors and the patterns of item relationships, is the same across different groups. It indicates whether the measurement tool captures the same theoretical structure across all groups, (2) Metric Invariance: Known as factorial weight invariance, this level is achieved when the factor loadings (weights) are the same across groups. Each item holds the same meaning for every group, (3) Scalar Invariance: This level is reached when the item intercepts (constants) are the same across groups. It ensures that the factor means can be reliably compared between groups, meaning that differences in means are not due to biases in the tool, (4) Residual Invariance: Also referred to as error or residual invariance, this occurs when the error terms (residuals) for items are the same across groups. This ensures that the discrepancies between observed and predicted values are equivalent across groups. Quality indices such as GFI, RMSEA, and Chi-square can be used to test MI [29]. In addition, Meredith [31] emphasizes that MI ensures that the scale operates the same way across different groups, such as demographic factors like gender or culture, and that MI is essential to ensure that differences between groups reflect true differences and not flaws in the scale. Vandenberg and Lance [16] and Dogruyol et al. [32] emphasize the importance of ensuring MI across its different levels.

7. PREVIOUS RESEARCH

The researcher reviewed a number of previous studies related to the topic of the study. Given the importance and wide use of the topic in research, several previous studies were found, including:

A study by [11] aimed to examine the attitudes of undergraduate students towards AI. The sample consisted of 460 university students (196 male and 264 female). The study used the GAAIS developed by [1], which was translated into Vietnamese. The results showed a Cronbach’s alpha value of 0.705, indicating acceptable reliability. The findings revealed that undergraduate students had moderately positive attitudes toward AI. The results also showed no statistically significant differences in attitudes based on gender. Chai et al. [5] focused on developing a scale to measure various factors that shape university students' behavioral intentions toward AI, including psychological attitudes. It involved 907 Chinese university students. The study found that reliability and confirmatory factor analyses indicated that the scale had acceptable reliability and validity. Lintner [22] conducted a systematic review of the quality of AI scales to assist researchers in choosing AI evaluation tools. The study reviewed 22 studies and validated 16 scales targeting different social groups, including university students. In general, the scales showed good construct validity and internal consistency. The study also indicated that no scale had been tested for measurement invariance across cultures.

Wang and Chuang [33] emphasized the importance of understanding users' subsequent behaviors toward AI technology. It aimed to develop a scale for self-efficacy toward AI by reviewing previous literature to generate preliminary items, extract factors, and conduct confirmatory factor analysis to verify the scale’s structure. The scale was applied to a sample of 314 cases. The study concluded that the scale consisted of four factors and 22 items, with good fit, reliability, and validity. Cai et al. [6] focused on the validity and reliability of the Chinese version of the AI Threats Scale for Chinese adults and its correlation with the Positive and Negative Effects of AI Scale. The results of the study showed that exploratory factor analysis and confirmatory factor analysis confirmed the suitability of the model. Moreover, a correlation was found with the Positive and Negative Effects Scale.

In a study by [17] on the possibility of prediction through correlations between variables, the study indicates the use of confirmatory factor analysis for the scale and confirms that the factor structure of the scale consists of two factors: a positive factor and a negative one. By comparing the scale with another related scale, the validity of the scale was confirmed. The study also used the Big Five personality traits to predict attitudes toward AI technology. The results showed that introverts had more positive attitudes, while higher distrust in companies led to negative attitudes toward this technology, and higher general trust led to positive attitudes. Kaya et al. [18] examined personality traits, anxiety, and demographic variables and their impact on attitudes toward AI technology. The study also translated the GAAIS into Turkish. The sample consisted of 340 males and females. The study used a personality scale, an anxiety scale for AI technology, and a scale for attitudes toward AI. The results showed that the scale for attitudes toward AI achieved good construct validity and reliability. The results also indicated a prediction of positive attitudes, with the coefficients being ($b=-0.172$, $p=0.004$), and negative attitudes with coefficients ($b=0.120$, $p=0.019$). The prediction coefficients for anxiety regarding the formation AI were ($b=-0.379$, $p<0.001$), while the coefficients for anxiety about learning AI were ($b=-0.211$, $p<0.001$). The study's results highlighted that personality traits and anxiety were significant factors determining attitudes toward this technology.

Kieslich et al. [20] proposed a scale to measure the perceived threats of AI that could be applied across various fields. It was applied to a sample of 891 cases, and the study concluded that the scale exhibited internal consistency and construct validity for the statistical indicators. In a study conducted by [1], the GAAIS was developed and underwent preliminary statistical validation through exploratory factor analysis, which identified positive and negative subscales. Both subscales captured emotions in line with their respective constructs. The positive subscale reflected societal and personal benefits, while the negative subscale reflected concerns and negative emotions. The scale showed good psychometric indicators, with convergent and discriminant validity compared to existing scales.

Summaries of tasks completed by AI applications were obtained from news articles to validate general attitudes with specific applications of AI. These were categorized in terms of comfort and perceived capability. Comfort with specific applications was a strong indicator of general attitudes, as measured by the GAAIS, while perceived capability was a weaker indicator. Participants viewed AI applications involving big data (e.g., astronomy, law, and pharmacy) positively but viewed applications involving human judgment (e.g., medical treatment and psychological counseling) negatively. Applications with strong ethical implications caused greater discomfort than their perceived capabilities suggested. Survey data indicated that people have diverse views on AI.

In a study by [34] to verify the MI in the Cattell Intelligence Test based on gender, the study was applied to a sample of 681 male and female students. The researchers used confirmatory factor analysis to verify the FS of the scale. The results showed a good model fit to the data, with suitable fit indices: CFI=0.95, TLI=0.94, RMSEA=0.05, and SRMR=0.04, indicating the validity of the FS of the scale. The study results also confirmed MI across gender, making the scale suitable for fair comparison between groups. Petric and Atanasova [12] conducted a study to validate an expanded scale for measuring e-health literacy through construct validity, content validity, and measurement reliability. The results showed a good fit with the data, with values of ($\chi2=2508$, df=282, RMSEA=0.064, SRMR=0.070, CFI=0.90). The scale also exhibited good internal consistency ($\alpha=0.89$), and the analysis confirmed measurement reliability across gender by validating the structural, metric, and scalar invariance within the proposed limits. Sahin and Yildrim [35] reviewed studies on public attitudes toward AI and examined 132 studies based on inclusion criteria. The results showed Cronbach's alpha values for the overall scale and the positive and negative subscales: 0.881, 0.828, and 0.863, respectively, while McDonald's omega coefficients were 0.873 and 0.923 for the negative and positive subscales, respectively.

In a study [36], a measurement tool was developed to assess ethical behavior among university students and the validity of the scale was checked. The results indicated that the measurement tool achieved construct and content validity, and the results confirmed measurement equivalence across gender through the coefficients for structural and metric invariance within acceptable values. Weyland et al. [37] conducted a study to verify the validity and reliability of the Physical Activity Enjoyment Scale in English. The results showed an Omega coefficient of 0.95 for reliability. Confirmatory factor analysis supported a unidimensional structure and model fit, and the results also supported measurement stability across gender. Morales-Garcia et al. [38] focused on the psychometric properties of the attitude toward AI scale. The results of confirmatory factor analysis revealed a unidimensional factor structure with excellent model fit, with fit indices of ($\chi 2$=0.41, p=0.522, CFI=1.00, TLI=1.00, RMSEA=0.00, SRMR=0.00). The results also showed good internal consistency (α=0.94, ω=0.91), and tests of measurement invariance from structural to strict invariance confirmed that the scale was stable across gender.

Grassini [39] verified the FS and MI of a brief scale for measuring attitudes toward AI technology. The study was applied to a sample of 500 individuals. The results indicated favorable fit indices for the confirmatory factor analysis with CFI=0.98, TLI=0.97, SRMR=0.03, and RMSEA=0.04. The reliability coefficient was 0.85, and the study concluded that the scale is suitable for future studies. In a study [40] to verify the FS of the Barkley Executive Function Disorder Scale, the study was applied to a sample of 678 male and female students. Confirmatory factor analysis was used to verify the FS of the scale, and the fit indices were CFI=0.94, TLI=0.93, RMSEA=0.046, and SRMR=0.38, indicating a good model fit to the data. The results also confirmed MI across gender and specialization through different equivalence levels. The study concluded that the scale has a good FS, achieves MI, and can be used in studies and research with different samples.

In a study conducted by [41], the factor analysis method was used to assess the Mobile Phone Addiction Scale, which was applied to a sample of 306 high school students in Kuwait. The results of the study indicated a good fit between the model and the data, with high indicator values such as IFI = 0.912, GFI = 0.05, RMSEA = 0.056, TLI = 0.898, and CFI = 0.902. The scale also demonstrated high reliability. Shalabi's [42] study aimed to verify the FS and MI of the standardized Progressive Matrices test. The study used Raven's Progressive Matrices test, known for its ability to measure general intelligence through the analysis of visual patterns. The sample consisted of 1000 male and female students from middle and high school levels. The results confirmed that Raven's test is based on a single-factor model (general intelligence factor). The model fit indicators were excellent: CFI = 0.96, indicating excellent fit, TLI = 0.95, which is very good, RMSEA = 0.045, an excellent value (< 0.05), and SRMR = 0.03, also excellent (< 0.05). The study also verified MI across different levels. For configural invariance, the model showed structural equivalence between the two groups with CFI = 0.95 and RMSEA = 0.048. For metric invariance, the factor weights were equal between the two groups, and the Chi-Square Difference was not statistically significant ($p > 0.05$). For scalar invariance, the fixed values were equal between the two groups, with CFI = 0.94 and RMSEA = 0.049. For residual invariance, the standard errors were equal between the two groups, and the Chi-Square Difference was not statistically significant ($p > 0.05$). The results of the study indicated that the FS and MI were present in the scale.

Through reviewing the previous studies, the researcher found that some studies focused on studying the psychometric properties of scales in different environments as well as measurement invariance. However, the scale of general attitudes toward AI has not been studied in the Saudi environment, nor has measurement invariance been examined. Additionally, this study differs from most previous studies in terms of the sample, which represents a sample of undergraduate university students. The current study aligns with previous studies in terms of the methodology and some statistical methods used, but it differs in terms of the sample and the nature of the study.

III. METHOD

1. RESEARCH ETHICS

The study was approved by the Research Ethics Committee at Taif University with the number (HAO-02-T-105) and the date of 15/1/2024, as the British Psychological Society (2014) (second edition) indicates the necessity of complying with the Code of Ethics for Human Research issued by the society.

2. STUDY DESIGN

The researcher used the descriptive methodology, which is suitable for the nature of the current study.

3. STUDY POPULATION AND SAMPLE

The study population includes students from Taif University. The sample consisted of 461 students, as shown in Table 1, with 220 male students (47.7%) and 241 female students (52.3%). The average age of the sample was 20.58 years with a standard deviation of 1.69. The study tools were applied to the sample of students during the second semester of the 2023/2024 academic year.

Table 1. Sample description.

Gender	N	Mean	Std. Deviation
Male	220	20.19	1.64
Female	241	20.93	1.67
Total	461	20.58	1.69

4. STUDY TOOL

The study tool was the GAAIS developed by [1]. The scale consists of 20 items distributed across two dimensions (subscales): a positive subscale and a negative subscale. Each subscale reflects emotions in alignment with their value. The positive subscale includes opportunities, benefits, community and personal advantages, and positive emotions) and consists of 12 items. The negative subscale includes fears and negative emotions) and consists of eight items. The items are rated using a five-point Likert scale (Strongly Agree, Agree, Neutral, Disagree, Strongly Disagree). Responses to the positive items are scored from 1 to 5 (from Strongly Disagree to Agree), while for the negative items, the scoring is reversed. The scale developers recommend summing the values for each subscale separately, rather than combining them to form a total score for the entire scale.

In the current study, the researcher translated the scale from English to Arabic and then presented it, after translation, to a group of four specialists in the English language with professional expertise in the field, who reviewed both the original and the translated versions. Some items of the scale were modified based on their feedback. The revised version was then presented to 11 faculty members as judges to ensure the clarity, linguistic accuracy, and relevance of the items to their respective dimensions. Based on the feedback from the judges, the wording of some items was revised. The final version of the scale, which consisted of 20 items distributed across two dimensions, was thus reached, maintaining the original structure of the scale:

IV. RESULTS AND DISCUSSION

For data analysis, IBM SPSS v.27 was used to calculate the descriptive statistics of the sample, including means, standard deviations, and percentages. The Cronbach's alpha coefficient was also calculated to assess reliability. Additionally, JASP 0.18.1 software was used to perform Confirmatory Factor Analysis (CFA) and MI for the GAAIS. Through conducting Confirmatory Factor Analysis (CFA) on the data, the results regarding the GFI indicators showed the following: Table 2 presents the GFI results, which indicate a chi-square value of (611.77) with degrees of freedom (df = 169). This is lower than the chi-square value in the baseline model (3236.3), suggesting a significant improvement in the model's fit to the data after modifying the factors. The statistical significance value ($p < 0.001$) indicates significant statistical differences.

Al-Ajmi's [41] study points out that the chi-square value is affected by the degrees of freedom and is one of the main drawbacks of this indicator, as larger sample sizes may lead to the rejection of the model even if it is a good or close match to the proposed model., Bollen [43] proposed using the ratio of chi-square value to degrees of freedom as a fitting indicator to address the issue of chi-square value sensitivity. For an acceptable fit, the ratio should be less than five. By dividing the chi-square value by the degrees of freedom (169), the ratio was found to be (3.62), below five, indicating that the model is acceptable.

Table 2 indicates that the Comparative Fit Index (CFI) was (0.855), a high value that falls within the acceptable range. The GFI value increases as it approaches one. This index is considered one of the best comparison indicators [26], and the value indicates the acceptance of the model, as it falls within the acceptable range for model acceptance. The TLI value was 0.837, which falls within the acceptable range (in some studies, 0.80 ≤, [44]. Also, the Non-Normed Fit Index (NNFI) value was also 0.837, and these values fall within the acceptable range for the model 0.80 ≤, [44].

The RMSEA value was 0.075, which is below the commonly accepted threshold of 0.08 [45], indicating model fit. Additionally, the GFI value was 0.983, which further supports the model's good fit to the data and indicates excellent model fit [46]. Bollen's Incremental Fit Index (IFI) was also used, with a value of (0.856), and the Relative Noncentrality Index (RNI) with a value of (0.855). These are indicators within the acceptable range and confirm the model's fit to the data. In general, the model can be accepted based on the commonly recognized index criteria, meaning that the model used is applicable and explains the data well.

Table 2. Fit indices.

Index	Value
χ^2 (df =190) Baseline Model	3236.33
χ^2 (df = 169, p < 0.001) *Factor Model*	611.77
CFI	0.855
TLI	0.837
NNFI	0.837
IFI	0.856
RNI	0.855
RMSEA	0.075
GFI	0.980

Table 3 shows that the values of the standardized factor loadings range between 0.482 and 0.859, indicating the extent to which the items are saturated with their respective dimensions. The first factor (positive subscale) has loadings ranging from 0.412 to 0.784, which are considered acceptable to good, indicating that the items represent the dimension (positive subscale). The loadings for the second factor (concerns and negative emotions) range from 0.550 to 0.859, which are good values, making them representative of the second factor. This is further confirmed by the fact that all items are statistically significant (P < 0.001), indicating that all items are reliably associated with their corresponding factor, which strengthens the model's reliability. Additionally, the standard errors range from 0.035 to 0.058, which are very small values, meaning that the estimates are stable. Therefore, the model shows good to excellent factor loadings for most items.

Table 3. Factor loadings.

Factor loadings							
						95% Confidence Interval	
Factor	Indicator	Estimate	Std. Error	z-value	p	Lower	Upper
Factor 1	V1	0.694	0.058	12.005	< .001	0.581	0.808
	V2	0.595	0.049	12.155	< .001	0.499	0.691
	V4	0.660	0.043	15.433	< .001	0.576	0.744
	V5	0.565	0.040	14.130	< .001	0.486	0.643
	V7	0.763	0.049	15.503	< .001	0.666	0.859
	V11	0.513	0.036	14.303	< .001	0.442	0.583
	V12	0.480	0.041	11.630	< .001	0.399	0.560

Factor loadings							
						95% Confidence Interval	
Factor	Indicator	Estimate	Std. Error	z-value	p	Lower	Upper
	V13	0.670	0.055	12.294	< .001	0.563	0.777
	V14	0.412	0.033	12.640	< .001	0.348	0.476
	V16	0.600	0.052	11.626	< .001	0.499	0.701
	V17	0.591	0.040	14.653	< .001	0.512	0.670
	V18	0.784	0.047	16.610	< .001	0.692	0.877
Factor 2	v3_RE	0.626	0.051	12.240	< .001	0.526	0.726
	V6_RE	0.550	0.043	12.816	< .001	0.466	0.635
	V8_RE	0.772	0.045	17.164	< .001	0.684	0.861
	V9_RE	0.641	0.057	11.249	< .001	0.529	0.752
	V10_RE	0.859	0.049	17.688	< .001	0.763	0.954
	V15_RE	0.734	0.050	14.734	< .001	0.636	0.832
	V19_RE	0.607	0.054	11.204	< .001	0.500	0.713
	V20_RE	0.705	0.055	12.712	< .001	0.596	0.813

As shown in Table 4, the values of the Average Variance Extracted (AVE) for the dimensions of the scale were 0.367 and 0.385. These are values that can be accepted, as Fornell and Larcker [47] indicated that an AVE less than 0.5 can be accepted if the composite reliability (ω) is greater than 0.6. In this study, we find that the composite reliability (ω) for the subscales is 0.869 and 0.832, respectively, and for the overall scale, it is 0.885.

Table 4. Average variance extracted.

Factor	AVE
Factor1	0.367
Factor2	0.385

In addition, the scale's reliability was verified as shown in Table 5, by calculating Cronbach's alpha reliability coefficient for the first subscale, which was 0.869. This is a high value for the reliability coefficient. Moreover, the Omega reliability coefficient for the first subscale was also 0.869, which is a high value and confirms the reliability of this subscale. The Cronbach's alpha reliability coefficient for the second subscale was calculated to be 0.832, and the Omega reliability coefficient for the second subscale was 0.828. These are high values exceeding the required threshold of 0.70 [46], indicating that the scale demonstrates good reliability.

Table 5. Reliability

Factor	Coefficient ω	Coefficient α
Factor 1	0.869	0.869
Factor 2	0.832	0.828
Total	0.885	0.803

QUBAHAN ACADEMIC JOURNAL
VOL. 5, NO. 2, April 2025
https://doi.org/10.48161/qaj.v5n2a1672

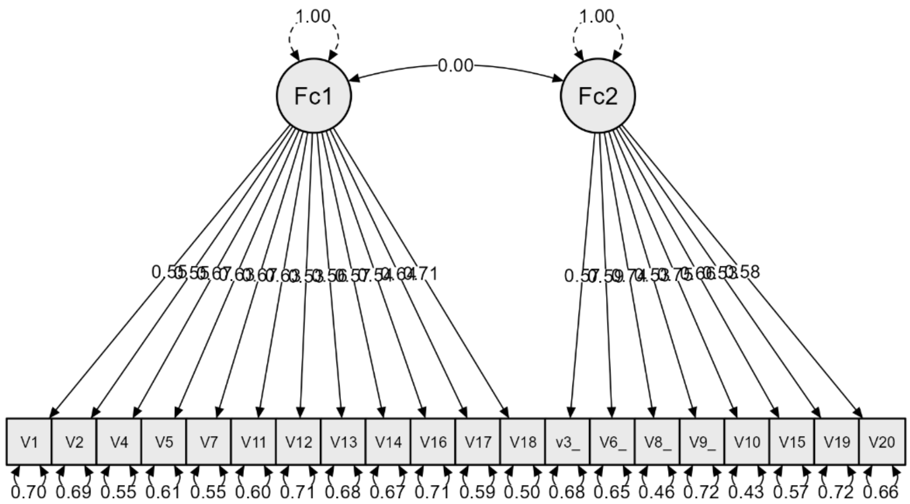

FIGURE 1. Standardized factor loadings for the GAAIS.

Figure 1 shows the Confirmatory Factor Analysis (CFA) model for measuring general attitudes towards AI, consisting of two subscales. The first factor, FC1, represents the positive subscale and the benefits and usefulness, while the second factor, FC2, represents the negative subscale and the fears and negative emotions. The correlation value between the two factors is (0.00), indicating that there is no correlation between the two subscales, which strengthens the idea that each subscale is distinct from the other. To confirm the measurement invariance of the GAAIS, a multi-group confirmatory factor analysis was used. Four types of measurement invariance were checked: A multi-group confirmatory factor analysis was conducted. Four types of MI were tested to ensure the MI of the GAAIS:

1. CONFIGURAL INVARIANCE

According to Table 6, the chi-square value is 848.901 with 338 degrees of freedom. When calculating the ratio of the chi-square value to the degrees of freedom, the value is 2.51, which is less than five and is considered an acceptable value [43]. Additionally, the GFI is 0.978, a high and acceptable value. The CFI is 0.839, and the TLI is 0.819, both within the acceptable range. The RMSEA is 0.08, which also falls within the acceptable range. These indices indicate that configural invariance is achieved, meaning that participants from both groups (male and female) understand the items in the same way [47]. This result means that the underlying factor structure of the model is stable across groups.

Table 6. Fit indices for configural invariance.

Index	Value
χ^2 $(df = 338, \text{p} < 0.001)$	848.901
RMSEA	0.081
GFI	0.978
CFI	0.839
TLI	0.819

2. METRIC INVARIANCE

According to Table 7, the chi-square value is 863.381 with 356 degrees of freedom. When calculating the ratio of the chi-square value to the degrees of freedom, the value is 2.43, which is less than five and is considered an acceptable value [43].

Furthermore, the GFI is 0.978, a high and excellent value to the model. The CFI is 0.840, and the TLI is 0.829, both within the acceptable range. The RMSEA value of 0.08 is considered acceptable as it falls within the fit criteria, indicating the adequacy of the model. This suggests that the scale demonstrates MI, making it suitable for use across different groups [29].

Table 7. Fit indices for Metric Invariance.

Index	Value
χ^2 (df = 356, p < 0.001)	863.381
RMSEA	0.079
GFI	0.978
CFI	0.840
TLI	0.829

3. SCALAR INVARIANCE

According to Table 8, the Chi-square value is 930.892 with 374 degrees of freedom. When calculating the ratio of the chi-square value to the degrees of freedom, the value is 2.48, which is less than five and is considered an acceptable value [43].

Table 8. Fit indices for scalar invariance.

Index	Value
χ^2 (df = 374, p < 0.001)	930.82
RMSEA	0.080
GFI	0.976
CFI	0.824
TLI	0.821

Additionally, the GFI is 0.976, a high and acceptable value. The CFI is 0.824, and the TLI is 0.821, both within the acceptable range. The RMSEA is 0.08, which also falls within the acceptable range. These indices indicate that scalar invariance is achieved, meaning that the intercepts of the items on the scale are equal across the two groups in the study. This suggests that the scale can be used to compare the latent variable means between the groups [29].

4. RESIDUAL INVARIANCE

According to Table 9, the chi-square value is 978.987 with 394 degrees of freedom. When calculating the ratio of the chi-square value to the degrees of freedom, the value is 2.48, less than five, and is considered an acceptable value [43]. Additionally, the GFI is 0.975, a high and acceptable value. The CFI is 0.815, and the TLI is 0.829, both within the acceptable range. The RMSEA is 0.08, which also falls within the acceptable range. These indices indicate that residual invariance is achieved.

Table 9. Fit indices for residual invariance.

Index	Value
χ^2 (df = 394, p < 0.001)	978.987
RMSEA	0.080
GFI	0.975
CFI	0.815
TLI	0.822

The evaluation of public attitudes towards AI is crucial due to the rapid development of this technology and its impact on university students. To the best of the researcher's knowledge, no study has been conducted to explore university students' attitudes towards AI in the local context. Therefore, this research is one of the first studies on the attitudes of undergraduate students in Saudi Arabia towards AI. Regarding the factor structure and reliability of the scale.

The results showed that the GAAIS fit the confirmatory factor analysis model, with the following GFI indices: $\chi^2/df=611.77/169=3.62$), CFI = 0.855, TLI = 0.837, and RMSEA = 0.075. These values confirm the construct validity and the scale's fit to the data. The Composite Reliability coefficients were 0.828 and 0.869, while the Cronbach's alpha coefficients were 0.832 and 0.869 for the positive subscale (opportunities, benefits, social and personal utility, and positive emotions) and the negative subscale (concerns and negative emotions), respectively. These values are somewhat similar to the original scale [1], which confirms the validity and reliability of the scale for measuring attitudes towards AI.

The study results also indicated the presence of configural, metric, scalar, and residual invariance between the male and female samples. It was found that the goodness-of-fit indices fall within the range proposed by [7]. The results of the current study align with the findings of several studies, including those by [1, 17, 20, 38]. These studies demonstrated that the GAAIS possesses construct validity according to confirmatory factor analysis and serves as an effective tool for measuring general attitudes toward AI.

Additionally, the results of this study align with the findings of [12, 34, 37-38, 40, 42]. All of which demonstrated the verification of the factor structure through fit indices and confirmed the MI across different groups through various levels of equivalence. These consistent findings support the reliability and validity of the scale across different contexts and groups.

V. CONCLUSION AND RECOMMENDATIONS

Understanding university students' attitudes toward AI is crucial for educators, policymakers, and professionals aiming to address concerns, manage negative emotions, and enhance the learning process. It is essential for preparing students for a future where AI plays a prominent and important role. This can be achieved by providing an appropriate scale tailored to the local context. The results of the current study show that construct validity for the scale was achieved. This validates the scale's applicability to university students based on gender (male, female), allowing researchers to use it in studies where attitudes toward AI are among the variables. When using the GAAIS to compare male and female students' performance, there are no concerns regarding interpreting whether any performance differences are due to gender, based on the results obtained from the sample, as the MI confirms the fairness of the scale. There are some limitations to the current study. First, the sample was limited to undergraduate students and did not include other academic levels, which resulted in a limited representation of the student population. Second, the study relied solely on a quantitative approach, and combining both quantitative and qualitative methods could contribute to a better understanding of university students' attitudes toward AI. In light of the study findings, the study recommends applying the GAAIS to other samples from different educational and non-educational communities. In addition, researchers can use the GAAIS alongside other variables such as personality, motivation, anxiety, and other psychological variables. Moreover, the GAAIS can be applied to other samples from different study populations.

Funding Statement

There was not any external funding for this study.

Author Contributions

The author alone carried out this investigation. The author was in charge of the study's conception, methodological design, data collection and analysis, result interpretation, and article preparation and revision.

Conflicts of Interest

No conflicts of interest are disclosed by the author.

Data Availability Statement

The author may provide data upon request.

Acknowledgments

The author would like to thank the editor and reviewers for their help in getting the article prepared for publication.

REFERENCES

1. Schepman, A., & Rodway, P. (**2020**). Initial validation of the General Attitudes towards Artificial Intelligence Scale. *Computers in Human Behavior Reports*. https://doi.org/10.1016/j.chbr.2020.100014
2. Russell, S., & Norvig, P. (**2021**). *Artificial intelligence: A modern approach* (4th ed.). Pearson.
3. Verma, M. (**2018**). Artificial intelligence: Its scope in different areas with special reference to the field of education. *Online Submission*, 3(1), 5–10.
4. Schepers, J., & Wetzels, M. (**2007**). A meta-analysis of the technology acceptance model: Investigating subjective norm and moderation effects. *Information & Management*, 44(1), 90–103. https://doi.org/10.1016/j.im.2006.10.007
5. Chai, C., King, R., & Zho, Y. (**2024**). Development and validation of the Artificial Intelligence Learning Intention Scale (AILIS) for university. *SAGE Open*, 14(2). https://doi.org/10.1177/21582440241242188
6. Cai, J., Xu, Z., Sun, X., Guo, X., & Fu, X. (**2023**). Validity and reliability of the Chinese version of Threats of Artificial Intelligence Scale (TAI) in Chinese adults. *Psicologia: Reflexão e Crítica*, 36(5). https://doi.org/10.1186/s41155-023-00247-1
7. Chen, F. F., Sousa, K. H., & West, S. G. (**2005**). Testing measurement invariance of second-order factor models. *Structural Equation Modeling: A Multidisciplinary Journal*, 12(3), 471–492. https://doi.org/10.1207/s15328007sem1203_7
8. Venkatesh, V., & Davis, F. D. (**2003**). User acceptance of information technology: Toward a unified view. *MIS Quarterly*, 27(3), 425–478. https://doi.org/10.2307/30036540
9. Davis, F. D. (**1989**). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly*, 13(3), 319–340. https://doi.org/10.2307/249008
10. Stein, J.-P., Messingschlager, T., Gnambs, T., Hutmacher, F., & Appel, M. (**2024**). Attitudes towards AI: Measurement and associations with personality. *Scientific Reports*, 14, 2909. https://doi.org/10.1038/s41598-024-53335-2
11. Le, M. T. (**2024**). Exploring undergraduate students' general attitudes towards Artificial Intelligence: A perspective from Vietnam. *Journal of Language and Cultural Education*, 12(3). https://doi.org/10.2478/jolace-2024-0014
12. Petrič, G., & Atanasova, S. (**2024**). Validation of the extended e-health literacy scale: Structural validity, construct validity and measurement invariance. *BMC Public Health*, 24, 1991. https://doi.org/10.1186/s12889-024-19431-8
13. Protzko, J. (**2022**). Invariance: What does measurement invariance allow us to claim? *Educational and Psychological Measurement*. https://doi.org/10.1177/00131644241282982
14. Clawson, R. E., Bean, R. A., Dyer, W. J., Bradford, A. B., Anderson, S. R., & Lee, C.-T. P. (**2023**). Examining a key measure of youth disclosure to parents for measurement invariance across time and reporters. *Journal of Child and Family Studies*, 32, 1765–1775. https://doi.org/10.1007/s10826-022-02388-w
15. Kline, R. B. (**2015**). *Principles and practice of structural equation modeling* (3rd ed.). Guilford Press.
16. Vandenberg, R. J., & Lance, C. E. (**2000**). A review and synthesis of the measurement invariance literature: Suggestions, practices, and recommendations for organizational research. *Organizational Research Methods*, 3(1), 4–70. https://doi.org/10.1177/109442810031002.
17. Schepman, A., & Rodway, P. (**2023**). The General Attitudes towards Artificial Intelligence Scale (GAAIS): Confirmatory validation and associations with personality, corporate distrust, and general trust. *International Journal of Human–Computer Interaction*, 39(13), 2724–2741. https://doi.org/10.1080/10447318.2022.2085400
18. Kaya, F., Aydin, F., Schepman, A., Rodway, P., Yetişensoy, O., & Demir Kaya, M. (**2024**). The roles of personality traits, AI anxiety, and demographic factors in attitudes toward artificial intelligence. *International Journal of Human–Computer Interaction*, 40(2), 497–514. https://doi.org/10.1080/10447318.2022.2151730
19. Wei, J. (**2018**). Research progress and application of computer artificial intelligence technology. *MATEC Web of Conferences*, 176, 01043. https://doi.org/10.1051/matecconf/201817601043
20. Kieslich, K., Lünich, M., & Marcinkowski, F. (**2021**). The Threats of Artificial Intelligence Scale (TAI): Development, measurement and test over three application domains. *International Journal of Social Robotics*, 13, 1563–1577. https://doi.org/10.1007/s12369-020-00734-w

21. Saklaki, A., & Gardikiotis, A. (**2024**). Exploring Greek students' attitudes toward artificial intelligence: Relationships with AI ethics, media, and digital literacy. *Societies*, 14, 248. https://doi.org/10.3390/soc14120248
22. Lintner, T. (**2024**). A systematic review of AI literacy scales. *npj Science of Learning*, 9(50). https://doi.org/10.1038/s41539-024-00264-4
23. Lee, J. (**2018**). Human attitudes toward artificial intelligence: Understanding the role of perceived risks and benefits. *AI & Society*, 33(1), 1–10. https://doi.org/10.1007/s00146-017-0730-8
24. Sun, P., Zhang, Y., & Wang, Y. (**2020**). The evolution of public attitudes toward AI technologies: A longitudinal study. *Computers in Human Behavior*, 112, 106446. https://doi.org/10.1016/j.chb.2020.106446
25. Parasuraman, A. (**2000**). Technology Readiness Index (TRI): A multiple-item scale to measure readiness to embrace new technologies. *Journal of Service Research*, 2(4), 307–320. https://doi.org/10.1177/109467050024001
26. Kline, R. B. (**2016**). *Principles and practice of structural equation modeling* (4th ed.). Guilford Press.
27. Byrne, B. M. (**2012**). *Structural equation modeling with Mplus: Basic concepts, applications, and programming*. Routledge.
28. Brown, T. A. (**2015**). *Confirmatory factor analysis for applied research*. Guilford Press.
29. Cheung, G. W., & Rensvold, R. B. (**2002**). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling*, 9(2), 233–255. https://doi.org/10.1207/S15328007SEM0902_5
30. Ding, Z., Ng, F., & Wang, J. (**2014**). Testing trust scale measurement invariance in project teams. *Journal of Engineering, Design and Technology*, 12(2), 209–222. https://doi.org/10.1108/JEDT-04-2012-0017
31. Meredith, W. (**1993**). Measurement invariance, factor analysis, and factorial invariance. *Psychometrika*, 58(4), 525–543. https://doi.org/10.1007/BF02294825
32. Doğruyol, S., Uzun, N. B., Aygar, B. B., & Yücedağlar, A. (**2024**). Examining measurement invariance of different SWLS measurement models according to gender. *Journal on Educational Psychology*, 17(3). https://doi.org/10.26634/jpsy.17.3.19823
33. Wang, Y.-Y., & Chuang, Y.-W. (**2024**). Artificial intelligence self-efficacy: Scale development and validation. *Education and Information Technologies*, 29, 4785–4808. https://doi.org/10.1007/s10639-023-12015-w
34. Zaraa, N., & Al-Eidan, K. (**2024**). Measurement invariance in the Cattell intelligence test among adults based on gender using multigroup confirmatory factor analysis. *Journal of Arts for Psychological and Educational Studies*, 6(1), 9–45.
35. Şahin, M. G., & Yıldırım, Y. (**2024**). The general attitudes towards artificial intelligence (GAAIS): A meta-analytic reliability generalization study. *International Journal of Assessment Tools in Education*, 11(2), 303–319. https://doi.org/10.21449/ijate.1369023
36. Al-Nuaimi, M., & Al-Emran, M. (**2024**). Development and validation of ICT unethical behavior scale among undergraduate students. *Current Psychology*, 43, 8760–8776. https://doi.org/10.1007/s12144-023-05038-6
37. Weyland, S., Kaushal, N., Fritsch, J., Strauch, U., & Jekauc, D. (**2024**). Validation and invariance testing of the English short physical activity enjoyment scale. *PLOS ONE*, 19(11), e0313626. https://doi.org/10.1371/journal.pone.0313626
38. Morales-García, W. C., Sairitupa-Sanchez, L. Z., Morales-García, S. B., & Morales-García, M. (**2024**). Adaptation and psychometric properties of an Attitude toward Artificial Intelligence Scale (AIAS-4) among Peruvian nurses. *Behavioral Sciences*, 14(5), 437. https://doi.org/10.3390/bs14060437
39. 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(5), 100–120. https://doi.org/10.3389/fpsyg.2023.1191628
40. Atia, A. (**2020**). Factor structure of the Barkley executive function disorder scale and its measurement invariance based on item response theory and multigroup confirmatory factor analysis among university students. *Educational Sciences*, 28(4), 327–438.
41. Al-Ajmi, M. A. (**2020**). Confirmatory factor analysis of the smartphone addiction scale among a sample of high school students in Kuwait. *Scientific Journal of the Faculty of Education, Assiut University*, 20(5), 90–123.
42. Shalabi, S. (**2015**). Factor structure and measurement invariance of the Raven's progressive matrices test among middle and high school students based on the structural equation model. *Educational Sciences*, 23(4), 17–45.
43. Bollen, K. A. (**1989**). *Structural equation with latent variables*. John Wiley. https://doi.org/10.1002/9781118619179
44. Marsh, H. W., Hau, K. T., & Wen, Z. (**2004**). In search of golden rules: Comment on hypothesis-testing approaches to setting cutoff values for fit indexes. *Structural Equation Modeling*, 11(3), 320–341. https://doi.org/10.1207/s15328007sem1103_2
45. Browne, M. W., & Cudeck, R. (**1993**). Alternative ways of assessing model fit. *Sociological Methods & Research*, 21(2). https://doi.org/10.1177/0049124192021002005
46. Jöreskog, K. G., & Sörbom, D. (**1989**). *LISREL 7: A guide to the program and applications* (2nd ed.). SPSS Inc.
47. Fornell, C., & Larcker, D. F. (**1981**). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research*, 18(1), 39–50. https://doi.org/10.2307/3151312