Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Vietnamese Adaptation and Psychometric Evaluation of the AI Attitude Scale-4 Among University Students

Thi Duyen Vu
Van Hien University
Tuan Kiet Pham
Van Lang University
Thi Bao Ngan Nguyen
Van Hien University
An Dinh Duy Phan
Taipei Medical University
Minh Thuan Ly
Van Hien University
Huu Son Hai Le
Van Lang School of Technology, Van Lang University
Simone Grassini
simone.grassini@uib.no

University of Bergen

Article

Keywords: attitudes toward artificial intelligence, AI Attitude Scale, AIAS-4, psychometric validation, measurement invariance, Vietnamese university students

Posted Date: November 12th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11256339/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License. Read Full License

Additional Declarations: No competing interests reported.

Abstract

Attitudes toward artificial intelligence (AI) shape how students adopt AI, and the four-item AI Attitude Scale (AIAS-4) offers a brief measure of these attitudes. No validated Vietnamese version exists, and equivalence cannot be assumed from direct translation. We translated and culturally adapted the AIAS-4 with expert content validation, expert linguistic review, and a pilot test (n = 81), then evaluated it in 698 Vietnamese university students split into exploratory and confirmatory subsamples (n = 349 each), with a two-week retest in 60 participants. All items reached an item-level content validity index of 1.00. Exploratory and confirmatory factor analyses supported a single factor (loadings .82–.93; *CFI* = .99, *RMSEA* = .057). Internal consistency was high (α and ω = .92) and test–retest reliability was good (*ICC* = .84), with no mean change but individual-level proportional bias. Scores correlated modestly with eHealth literacy and attitudes toward seeking professional psychological help (r = .23–.37). Strict invariance held across age groups; invariance across gender was supported but less consistent at the metric level. The Vietnamese AIAS-4 is suitable for group-level research on AI attitudes in higher education, whereas small individual changes should be interpreted cautiously. Responsiveness and cross-linguistic equivalence remain to be established.

Introduction

Artificial intelligence (AI) is increasingly embedded in education, healthcare, and digital services, and university students are among its most frequent users through chatbots, generative tools, and recommendation systems[1, 2]. Students recognize the benefits of AI but also voice concerns about privacy, data security, and future employment[3, 4]. How users perceive and respond to AI shapes its effective use: attitudes toward AI are associated with willingness to use AI, technology acceptance, trust, and behavioral intentions[5, 6], which makes their measurement relevant to research, education, and implementation[7, 8].

An attitude is a psychological tendency expressed by evaluating an object favorably or unfavorably[9] and may comprise cognitive, affective, and behavioral components[10]. Attitude toward AI is accordingly a person's relatively general evaluation of AI and its potential benefits, effects, and consequences[11], although existing measures operationalize these components to different degrees[12]. It should be distinguished from related constructs such as AI literacy, AI anxiety, trust in AI, and intention to use AI[13–15], and from technology-acceptance models, which focus on the usefulness and ease of use of a specific system[16] rather than on AI as a class of technology and its wider implications[11, 14, 17]. Clearly defining what each instrument measures is therefore essential, because these concepts are not interchangeable[12, 18].

Because attitude toward AI is a latent construct, it should be assessed with validated multi-indicator measures rather than isolated questions[19]. When an instrument is adapted to a new language, semantic and conceptual equivalence and cultural relevance must be established[20] alongside structural validity, reliability[21], and measurement invariance across groups[22]. Standardized measures then allow researchers to identify correlates of AI attitudes, evaluate AI literacy programs, and compare learner groups[23].

Several instruments assess general attitudes toward AI. The General Attitudes towards Artificial Intelligence Scale (GAAIS) uses 20 items in positive and negative subscales[11]; the five-item Attitude Towards Artificial Intelligence scale (ATAI) distinguishes acceptance from fear and has been applied in German, Chinese, English, and Arabic[24, 25]; and the ATTARI-12 captures cognitive, affective, and behavioral expressions within a predominantly unidimensional structure[26]. The AI Attitude Scale (AIAS-4) is a four-item, single-factor measure of general positive attitudes toward AI that covers expected benefits for one's life and work, intention to use AI, and its perceived value for humanity[18]. Its unidimensional structure and reliability were supported in the development study and replicated in a large sample of Norwegian students and in two Turkish and one Spanish adaptation[27–30]. The brief format eases inclusion in large multi-measure surveys while assessing AI at a general rather than application-specific level[18].

In Vietnam, attitudes toward AI have been examined among healthcare students[31] and with a Vietnamese GAAIS in 460 students from five universities, which reported acceptable internal consistency[32]. These studies focused on descriptive statistics, reliability, and group comparisons rather than comprehensive psychometric evaluation, and the GAAIS differs from the AIAS-4 in structure and scope. To our knowledge, no study has translated, culturally adapted, and psychometrically evaluated the AIAS-4 in Vietnamese; because the scale was developed in English, semantic and conceptual equivalence cannot be assumed from direct translation[20].

The present study therefore translated and culturally adapted the AIAS-4 into Vietnamese and evaluated its psychometric properties among university students. We assessed face and content validity, tested whether the one-factor structure was retained in two independent subsamples, estimated internal consistency and two-week test–retest reliability, examined relations with external variables, and tested measurement invariance across prespecified age and gender groups. We hypothesized that the Vietnamese AIAS-4 would retain its one-factor structure, show adequate reliability and temporal stability, correlate meaningfully with the external variables, and be invariant across the demographic groups.

Methods

Study Design

A psychometric validation study comprising a cross-sectional main survey and a two-week prospective test-retest component, using self-report questionnaires among Vietnamese students.

Translation and Cross-Cultural Adaptation of the AIAS-4

The translation and cross-cultural adaptation of the AIAS-4 were guided by established recommendations for the adaptation of patient- and self-reported measurement instruments, including the practical framework proposed by Cruchinho et al.[33]. The procedure was designed to preserve the conceptual meaning of the original items while producing Vietnamese wording that was linguistically natural and culturally appropriate for the intended target population. An audit trail was maintained throughout the process to document discrepancies, revisions, and the rationale for each translation decision (Supplementary Fig. S1).

Permission and preparation

Permission to translate and validate the AIAS-4 for use in Vietnam was obtained from the original instrument developer before the adaptation process commenced. The research team reviewed the construct definition, item content, response format, scoring procedure, and available evidence regarding the psychometric properties of the original scale. Translation materials were then prepared, including the original English version, standardized translator instructions, item-level comparison forms, and documentation templates for recording translation decisions.

Translation, Synthesis, Back-Translation, and Harmonization

The AIAS-4 was translated from English into Vietnamese through a multistep process. First, two bilingual translators independently prepared the forward translations. One translator was informed about the construct and the purpose of the scale, whereas the other translator was not given detailed information about the underlying concept. This arrangement was intended to maintain conceptual accuracy while also producing wording that sounded natural and understandable in Vietnamese. The two translations were then reviewed item by item by the translators and members of the research team. Differences in wording, sentence structure, tone, and conceptual emphasis were discussed until a single synthesized version was agreed upon.

The synthesized Vietnamese version was subsequently translated back into English by two bilingual translators who had not participated in the forward-translation stage and had not seen the original items. The back-translated versions were compared with the original AIAS-4 to identify possible changes in meaning, omissions, or added interpretations. Any discrepancies were discussed and revised when needed. Finally, a harmonization meeting was held with the translators, research team, and relevant language and subject specialists. All versions were reviewed together with attention to semantic, idiomatic, experiential, and conceptual equivalence, as well as the appropriateness of the instructions and the 10-point response format for Vietnamese respondents. Revisions were made by consensus, resulting in the pre-final Vietnamese AIAS-4 used for expert evaluation.

Expert Review and Content Validity

Expert panel

The pre-final Vietnamese AIAS-4 was independently evaluated by eight experts with relevant backgrounds in psychology, psychometrics, research methodology, language, education, and/or artificial intelligence. Experts were selected based on their academic qualifications, professional experience, and familiarity with scale development or cross-cultural adaptation. Each expert received the original English items, the proposed Vietnamese translations, a description of the construct, standardized rating instructions, and space for qualitative comments and suggested revisions.

Rating procedure

Each item was assessed separately on three dimensions: relevance, clarity, cultural appropriateness and semantic equivalence. A four-point ordinal response scale was used: (1) = not appropriate or not clear; (2) = requires major revision; (3) = appropriate with minor revision; (4) = highly appropriate.

For the calculation of content-validity indices, ratings of 3 or 4 were classified as agreement that the item was content valid, whereas ratings of 1 or 2 were classified as nonagreement. Experts also provided written comments regarding wording, conceptual equivalence, and cultural suitability. Quantitative ratings and qualitative feedback were reviewed jointly before finalizing the wording for the linguistic review.

Content validity was evaluated separately for relevance, clarity, cultural appropriateness, and semantic equivalence. The *I-CVI* was calculated as the proportion of eight experts rating an item as 3 or 4, while *S-CVI/Ave* represented the mean *I-CVI* and *S-CVI/UA* represented universal agreement across items. Modified kappa was also calculated to adjust for chance agreement. An *I-CVI* of at least .78 and an *S-CVI/Ave* of .90 or higher were considered acceptable and excellent, respectively[34].

Expert linguistic review

The content-validated Vietnamese AIAS-4 was then reviewed by the same eight experts. Each expert completed the scale and was individually interviewed using verbal probing about item comprehensibility, key terms, wording clarity, cultural relevance, potential ambiguity, and use of the 10-point response format. Interview notes were reviewed for recurrent interpretation problems or wording that departed from the intended construct. Revisions were considered only for difficulties raised by several experts, without altering the conceptual content of the original items.

Pilot Testing

The Vietnamese AIAS-4 was pilot tested with 81 participants before the main survey. Preliminary internal consistency was assessed using Cronbach's *alpha*, with a coefficient of at least .70 considered acceptable. The pilot results did not lead to any changes to the scale, and all four items were retained for the main validation study.

Participants, Data Collection, and Ethical Considerations

Ethical approval

and informed consent

The study was approved by the Ethics Review Board of the University of Social Sciences and Humanities, Vietnam National University Ho Chi Minh City (ERB application No. *41-2026*; certificate No. *569/GXN-XHNV-ĐN&QLKH*, 5 May 2026). All procedures were performed in accordance with the Declaration of Helsinki and relevant institutional guidelines and regulations. Informed consent was obtained electronically from all participants before they accessed the questionnaire.

Study setting and recruitment

The survey was administered via Google Forms, a platform for building online surveys and managing datasets. Students accessed the survey via a link or QR code presented on flyers, which were widely shared on social media (e.g., Facebook and LinkedIn) and within the researcher's professional networks. There were two conditions: participants must be at least 18 years old and be undergraduate students at Vietnamese universities. If they met the screening criteria, they would be presented with the information statement, in which they would be informed of the research's goal and asked to indicate their consent to participate in the survey. Participants were asked to provide an email address solely for follow-up contact to complete the second survey for the 2-week test-retest reliability assessment. A small company with 60 subsamples of participants completed the questionnaire again two weeks later. Email addresses were stored securely, were accessible only to the research team, and were deleted after completion of the follow-up data collection. The survey was a self-report questionnaire, including demographic questions (e.g., age, gender, education) and related scales, and took approximately 10–15 minutes to complete. The first recruitment process took place between 17th May and 29th May 2026, and the second recruitment for test-retest reliability happened between 31st May and 14th June 2026. All datasets were downloaded from Google Forms and analyzed in de-identified form.

The main validation sample comprised 698 participants. The sample was divided into two non-overlapping subsamples of equal size using *the SOLOMON* procedure, with 349 participants assigned to exploratory factor analysis and 349 to confirmatory factor analysis. The sociodemographic characteristics of both subsamples are reported in Table 1. A separate subsample of 60 participants completed the Vietnamese AIAS-4 again after a 2-week interval to assess test-retest reliability and agreement.

Table 1

Sociodemographic Characteristics of the EFA and CFA Subsamples

Variable	Category	EFA	CFA
n_{total} = 698		Frequency (%)	Frequency (%)
		349 (50)	349 (50)
Gender	Male	93 (26.6)	99 (28.4)
	Female	256 (73.4)	250 (71.6)
Age	18–20	301 (86.2)	271 (77.7)
	21 and above	48 (13.8)	78 (22.3)
Field of majors	Education	5 (1.4)	8 (2.3)
	Social Sciences	125 (35.8)	138 (39.5)
	Business and Tourism	9 (2.6)	7 (2.0)
	Natural Sciences	7 (2.0)	8 (2.3)
	Engineering	4 (1.1)	2 (.6)
	Other	199 (57.0)	186 (53.3)
Religion	No religion	149 (42.7)	149 (42.7)
	Practicing religion	72 (20.6)	76 (21.8)
	Non-practicing believer	128 (36.7)	124 (35.5)
Living Arrangement	Family/relatives	163 (46.7)	159 (45.6)
	Dormitory	9 (2.6)	9 (2.6)
	Roommates/friends	121 (34.7)	124 (35.5)
	Living alone	53 (15.2)	53 (15.2)
	Other	3 (.9)	4 (1.1)
Education in Mental health	No	211 (60.5)	209 (59.9)
	Yes	138 (39.5)	140 (40.1)

Note. EFA = exploratory factor analysis; CFA = confirmatory factor analysis. Percentages were calculated within each subsample. Totals may not equal 100% because of rounding.

Measures

Target instrument AIAS-4

The Artificial Intelligence toward Attitude Scale-4 (AIAS-4) is a brief self-report measure developed by Grassini[18] to assess general attitudes toward artificial intelligence. The scale consists of four items reflecting perceived positive impacts of AI on life, work, future technology use, and society. Items are rated from 1 (*completely disagree*) to 10 (*completely agree*). The overall score is calculated as the mean of the four items, ranging from 1 to 10, with higher scores indicating a more positive attitude toward AI. The original study supported a unidimensional structure and reported good internal consistency Cronbach's α = 0.90; McDonald's ω = 0.90[18]. As no validated Vietnamese version was available, the

AIAS-4 was translated, culturally adapted, and psychometrically evaluated in the present study. The original English items and their final Vietnamese versions are presented in Supplementary Table S1.

External validation measures

The eHealth Literacy Scale (eHEALS) was used to assess perceived ability to find, evaluate, and apply online health information. The scale contains eight items rated on a 5-point Likert scale, with higher scores indicating greater perceived eHealth literacy[35]. In the present study, the eHEALS demonstrated excellent internal consistency (Cronbach's α = .97; McDonald's ω = .97).

The Attitudes Toward Seeking Professional Psychological Help Scale-Short Form (ATSPPH-SF) was included to assess attitudes toward professional psychological help-seeking with 10 items[36]. The present analyses supported two dimensions: openness to seeking help (ATSPPH-O) and perceived value of professional services (ATSPPH-V). Both subscales showed good internal consistency, with α = .91 and ω = .91 for ATSPPH-O, and α = .90 and ω = .91 for ATSPPH-V.

Statistical Analysis

All analyses were conducted in *R* version 4.6.0 within *RStudio*, using the '*psych*', '*lavaan*', '*semTools*', and '*ggplot2*' packages. The main sample was divided into two equivalent, nonoverlapping subsamples using *the SOLOMON* procedure: one for exploratory factor analysis (EFA) and the other for confirmatory factor analysis (CFA)[37]. Categorical variables were summarized using frequencies and percentages, whereas continuous variables and item responses were examined using means, standard deviations, response distributions, corrected item-total correlations, and inter-item correlations. EFA used principal axis factoring, the number of factors was determined by parallel analysis and scree-plot inspection[38, 39], and no rotation was applied because one factor was retained.

Confirmatory Factor Analysis (CFA) was subsequently performed in the independent confirmatory subsample using *lavaan*[40] with maximum likelihood estimation. The model fit was evaluated using recommended indices: *RMSEA* and *SRMR* (< .08) and *CFI* (> .90), following Hair et al.[41]. Reliability and construct validity were examined using Cronbach's *alpha*, McDonald's *omega* total, composite reliability, maximal reliability, and average variance extracted. Relations with external variables were assessed using correlation coefficients, while discriminant validity was evaluated using HTMT and comparisons between AVE and MSV. Measurement invariance across gender and age groups was tested sequentially at the configural, metric, scalar, and strict levels. Test-retest reliability following the reporting recommendations of Koo and Li[42] was assessed using a two-way mixed-effects, absolute-agreement, single-measure ICC, together with SEM, MDC_{95}, mean bias, and Bland-Altman limits of agreement[43, 44]. All tests were two-tailed with α = .05. AIAS-4 items were treated as continuous given the 10-point response format.

A large language model was used for English language editing, all content was reviewed and verified by the authors.

Results

Content validity, linguistic review and pilot reliability

Eight experts in psychology, psychometrics, research methodology, linguistics, education, and AI rated the four Vietnamese items for relevance, clarity, and cultural appropriateness/semantic equivalence. All 96 ratings were 3 or 4; every item-level content validity index (I-CVI) and modified kappa was 1.00, and both scale-level indices (S-CVI/Ave and S-CVI/UA) were 1.00 for each criterion (Supplementary Tables S2-S4). Mean ratings were 3.97 for relevance, 3.94 for clarity, and 3.88 for cultural/semantic equivalence; Items 1 and 4 received the lowest cultural/semantic means (M = 3.75). No wording changes were required. Linguistic review identified no comprehension or interpretation problems. In the pilot (n =

81), Cronbach's *α* was .90 and a single factor emerged (Supplementary Table S5), so all four items were retained unchanged for the main study.

Exploratory factor analysis

Table 2
Exploratory Factor Analysis of Attitude towards Artificial Intelligence Using

Indicator		Items Content	Loadings	Communalities
			Factor 1	
1	Item 1	I believe that AI will improve my life.	.93	.86
2	Item 2	I believe that AI will improve my work.	.93	.86
3	Item 3	I think I will use AI technology in the future.	.90	.81
4	Item 4	I think AI technology is positive for humanity.	.88	.77
Bartlett's χ^2 (df)			1174.38*** (6)	
KMO			.84	
Eigenvalue			3.31	
Percent of total variance explained (%)			82.87	
Cumulative percent of total variance explained (%)			82.87	
Note. ***. p < .001; KMO = Kaiser-Meyer-Olkin coefficient; df = degrees of freedom.				

In the EFA subsample (*n* = 349), the Kaiser–Meyer–Olkin value was .84 and Bartlett's test of sphericity was significant, $\chi^2(6) = 1174.38$, $p < .001$. Parallel analysis supported a single factor (Supplementary Fig. S2). Principal axis factoring yielded loadings of .88-.93 and communalities of .77-.86, with an eigenvalue of 3.31 explaining 82.87% of the variance (Table 2) and no rotation was applied.

Confirmatory factor analysis and item analysis

Table 3
Item-level Estimates for Confirmatory Factor Analysis

Construct	Indicators	B	β	SE	z-value	95% CI
Attitude towards Artificial Intelligence (AIAS)	AIa1	1.00	.85	—	—	—
	AIa2	1.05	.93	.16	6.56***	[.74, 1.36]
	AIa3	.95	.82	.09	10.56***	[.77, 1.13]
	AIa4	.98	.87	.12	8.17***	[.74, 1.22]
Note. ***p < .001; B = Unstandardized Loadings coefficient; β = Standardized Loadings coefficient; SE = Standard Error; CI = Confidence Interval.						

Item-level estimates from Confirmatory Factor Analysis (CFA) were examined to evaluate the relationship between the observed variables and the latent construct (see Table 3 and Fig. 1). The standardized factor loadings ranged from .82 to .93, with all items exceeding .8 and statistically significant ($p < .001$), indicating that the observed indicators reflect the hypothesized construct. Additionally, the CFA results indicated that the one-factor measurement model showed good fit with $\chi^2/df = 2.14$, *CFI* = .99, *TLI* = .98, *RMSEA* = .057 (90% CI [.004, .131]), and *SRMR* = .010[41].

Figure 1.

Path diagram of the AIAS one-factor

Item-level analyses were conducted to examine the psychometric performance and homogeneity of four items within the AIAS-4. Corrected item-total correlations (CITCs) and inter-item correlations were examined (Supplementary Table S6). The CITC values, ranging from .77 to .86, indicated that all four items of the scale were adequately associated with the overall scale, with all coefficients exceeding the common recommendation threshold of .3[45]. In addition, the inter-item correlations ranged from .73 to .86, indicating adequate associations among items. Based on the item-level analysis results, none of the items were removed from AIAS-4.

Reliability, convergent and discriminant validity

Table 4

Reliability, Convergent Validity, and Discriminant Validity of the Study Measures

Variable	AIAS-4	eHEALS	ATSPPH-O	ATSPPH-V
Number of items	**4**	8	5	5
Cronbach's *α*	**.92**	.97	.91	.90
Guttman's *λ-2*	**.92**	.96	.91	.90
McDonald's *ω*	**.92**	.97	.91	.91
CR	**.93**	.97	.91	.91
MaxR(H)	**.94**	.98	.92	.91
AVE	**.75**	.80	.67	.71
MSV	**.14**	.20	.24	.24

Note. α = Cronbach's alpha; λ-2 = Guttman's lambda-2; ω = McDonald's omega; CR = Composite Reliability; MaxR(H) = Maximal Reliability; AVE = Average Variance Extracted; MSV = Maximum Shared Variance.

Internal consistency reliability of AIAS-4 was examined using Cronbach's *alpha*, Guttman's *lambda-2*, McDonald's *omega*, Composite Reliability (CR), and Maximal Reliability (MaxR(H)) values. The reliability coefficients ranged from .92 to .94, demonstrating strong reliability of the scale (see Table 4). Additionally, Average Variance Extracted (AVE) and Maximum Shared Variance (MSV) indices were used to evaluate the convergent and discriminant validity of the measurement construct, respectively[46]. The AVE value for the AIAS-4 was .75, indicating adequate convergent validity. Discriminant validity was further assessed by comparing the AVE with MSV. The MSV index was lower than the AVE, supporting the discriminant validity of the AIAS-4 relative to other constructs.

Table 5

Correlations between Scale-level of Interest

Variable	M	SD	1	2	3	4
AIAS	6.60	2.14	—	.35***	.37***	.23***
eHEALS	3.76	.91		—	.45***	.29***
ATSPPH-O	3.18	.59			—	.49***
ATSPPH-V	2.74	.78				—

Note. ***. p < .001, AIAS = AI Attitude; eHEALS = eHealth Literacy; ATSPPH-O = Openness to Seeking Professional Psychological Help; ATSPPH-V = Value and Need in Seeking Professional Psychological Help; M = Mean; SD = Standard Deviation.

AIAS-4 scores correlated positively with eHealth literacy (r = .35), openness to seeking professional psychological help (ATSPPH-O; r = .37), and perceived value of such help (r = .23), all p < .001 (Table 5). These small-to-moderate associations indicate meaningful but limited overlap with the external constructs. All heterotrait–monotrait (HTMT) ratios between the AIAS-4 and the other scales were below .45, well under the .85 criterion[47] (Supplementary Table S7).

Measurement invariance across age and gender

Table 6
Measurement Invariance of the AIAS-4 Across Gender and Age Groups

Group	Model	χ^2	df	χ^2/df	CFI	GFI	TLI	RMSEA	90% CI	ΔCFI	ΔRMSEA	SRMR
	Single	**4.27**	**2**	**2.14**	**.99**	**.98**	**.98**	**.057**	**[.004, .131]**	**–**	**–**	**.010**
Age												
	Configural	3.94	4	.99	1.00	.98	1.00	.000	[.000, .099]	–	–	.009
	Metric	4.49	7	.64	1.00	.99	1.00	.000	[.000, .058]	.000	.000	.012
	Scalar	7.73	10	.77	1.00	.99	1.00	.000	[.000, .060]	.000	.000	.018
	Strict	8.21	14	.59	1.00	.98	1.00	.000	[.000, .011]	.000	.000	.023
Gender												
	Configural	5.09	4	1.27	.998	.97	.99	.054	[.003, .113]	–	–	.010
	Metric	11.44	7	1.64	.994	.98	.98	.072	[.008, .117]	− .004	.018	.041
	Scalar	14.33	10	1.43	.995	.97	.98	.055	[.004, .104]	.001	− .017	.041
	Strict	16.82	14	1.20	.996	.96	.98	.038	[.009, .071]	.001	− .017	.050

Note. Age groups were defined as: Group 18–20 and Group 21 and above. CFI = Comparative Fit Index; TLI = Tucker-Lewis Index; GFI = Goodness-of-Fit Index; RMSEA (90% CI) = Root Mean Square Error of Approximation with Confidence Interval; SRMR = Standardized Root Mean Square Residual; df = degrees of freedom.

Measurement invariance of the AIAS-4 (see Table 6) was evaluated across age and gender groups using sequential configural, metric, scalar, and strict CFA models, with changes in model-fit indices considered alongside overall model fit[48, 49].

For age groups, all four models showed favorable fit, with *CFI* and *TLI* remaining at 1.00 and *RMSEA* at .000. Changes in *CFI* and *RMSEA* were consistently .000, while *SRMR* increased slightly from .009 in the configural model to .023 in the strict model. The chi-square difference tests were nonsignificant at the metric (p = .908), scalar (p = .356), and strict (p = .975) levels, supporting measurement invariance across age groups through the strict level. For gender groups, the configural model showed good fit (*CFI* = .998, *RMSEA* = .054, *SRMR* = .010). At the metric level, *CFI* decreased slightly (*ΔCFI* = − .004), whereas *RMSEA* and *SRMR* increased (*ΔRMSEA* = .018, *ΔSRMR* = .031). The scalar and strict models showed improvements in *CFI* and *RMSEA*, although *SRMR* increased to .050 in the strict model. Chi-square difference tests were nonsignificant at the metric (p = .096), scalar (p = .409), and strict (p = .646) levels. Overall, the results

supported measurement invariance across age groups, whereas the evidence across gender was generally favorable but less consistent at the metric level.

Test-retest reliability and measurement error

Table 7

Test-Retest Reliability, Measurement Error, and Agreement of the Vietnamese AIAS-4

Domain	Statistic	Result
Descriptive statistics	Time 1, *M* (*SD*)	6.60 (1.81)
	Time 2, *M* (*SD*)	6.72 (1.54)
Internal consistency	Time 1, *α* [95% CI]	.89 [.84, .93]
	Time 1, *ω*□	.89
	Time 2, *α* [95% CI]	.85 [.79, .91]
	Time 2, *ω*□	.86
Relative reliability	ICC(A,1) [95% CI]	.84 [.74, .90]
Systematic change	Mean bias, T2 - T1 [95% CI]	.12 [-.13, .37]
	Paired-samples test	$t(59) = .97$, $p = .338$
	Cohen's d	.12
Measurement error	SEM	.68
	MDC_{95}	1.90
Bland-Altman agreement	Lower LoA [95% CI]	-1.78 [-2.21, -1.35]
	Upper LoA [95% CI]	2.02 [1.59, 2.45]
Magnitude-dependent agreement	Proportional-bias slope, *b* [95% CI]	− .172 [-.32, − .02]
	Test of proportional bias	$p = .026$

Note. n = 60. α = Cronbach's alpha; ω□ = McDonald's omega total; ICC = intraclass correlation coefficient; SEM = standard error of measurement; MDC_{95} = minimal detectable change at the 95% confidence level; LoA = limits of agreement. ICC estimates were based on a two-way mixed-effects, absolute-agreement, single-measure model.

Sixty participants completed the scale twice, two weeks apart (Table 7). Mean scores were 6.60 (*SD* = 1.81) at Time 1 and 6.72 (*SD* = 1.54) at Time 2, with *α* = .89, 95% CI [.84, .93] and *α* = .85, 95% CI [.79, .91], respectively. The two-way mixed-effects, absolute-agreement, single-measure ICC(A,1) was .84, 95% CI [.74, .90]. Mean bias was .12, 95% CI [− .13, .37], $t(59) = .97$, $p = .338$, $d = .12$; because paired differences departed from normality (Shapiro–Wilk $W = .935$, $p = .003$), a Wilcoxon signed-rank test confirmed the absence of systematic change ($V = 794$, $p = .130$; Supplementary Table S10). The standard error of measurement was .68 and the minimal detectable change (MDC_{95}) was 1.90 points. Bland–Altman limits of agreement were − 1.78, 95% CI [− 2.21, − 1.35] and 2.02, 95% CI [1.59, 2.45], and the proportional-bias slope was negative, $b = -.172$, 95% CI [− .32, − .02], $p = .026$ (Fig. 2). The scale therefore ranked participants consistently and showed no mean shift, but individual-level agreement varied across the score range (Table 7; Supplementary Tables S8–S10 and Fig. S3).

Figure 2.

Bland-Altman Plot Assessing Test-Retest Agreement of the Vietnamese AIAS-4

The solid horizontal line indicates the mean bias, and the dashed lines represent the 95% limits of agreement. The regression line shows the presence of proportional bias between the paired differences and the mean scores.

Discussion

In this study, we translated, culturally adapted, and evaluated the psychometric properties of the AIAS-4 in the Vietnamese context. Overall, the findings support the Vietnamese version, particularly through the single-factor solution identified by exploratory factor analysis (EFA), its confirmation in an independent sample using confirmatory factor analysis (CFA), and the high factor loadings. These findings support the unidimensional structure reported in the original development study by Grassini[18]. The results for internal consistency, test-retest reliability, and measurement invariance were also generally favorable. However, the evidence regarding relations with external variables was modest, measurement invariance was not uniformly supported across all fit indices, and proportional bias in the retest analysis indicated the need for caution when interpreting changes in individual scores. The Vietnamese AIAS-4 should therefore be understood as a measure of general positive attitudes toward AI rather than a comprehensive assessment of all dimensions of AI attitudes.

Content validity evidence was strong: all expert ratings were 3 or 4, and the I-CVI, S-CVI, and modified kappa values met the criteria recommended by Polit et al.[34]. Items 1 and 4 nonetheless received slightly lower cultural/semantic ratings, probably because statements about AI “improving one's life” and “benefiting humanity” are broad and open to varied interpretation. High quantitative agreement supports content appropriateness but does not guarantee that every nuance of meaning survives translation; as other adaptations illustrate[30], equivalence must be secured through the whole adaptation process rather than literal translation alone[50].

With respect to structural validity, the factor-analytic findings supported the proposed one-factor structure of the Vietnamese AIAS-4 across two independent subsamples. In the EFA, parallel analysis supported the retention of a single factor, with strong factor loadings across the four items. CFA in an independent subsample provided further support for this structure, with high standardized factor loadings and good overall model fit. Examining the structure in separate subsamples helps reduce the risk of capitalizing on chance and provides additional evidence that the one-factor solution can be reproduced in another sample[51, 52]. These findings are consistent with the original AIAS-4 study by Grassini[18] and with subsequent adaptations among university students in Turkey[27] and Spain[28].

Internal consistency and model-based reliability were high and estimates from different coefficients fell within a narrow range, consistent with the original study[18] and the Turkish and Spanish adaptations[27, 28]. Convergence across estimators indicates that the reliability findings are robust to the choice of coefficient, which is more informative than interpreting Cronbach's alpha alone[53]. High internal consistency shows that the items cohere, not that the scale covers every facet of AI attitudes.

The AIAS-4 showed small to moderate positive correlations with eHealth literacy (eHEALS, $r = .35$, $p < .001$) and both dimensions of attitudes toward seeking professional psychological help (ATSPPH-O, $r = .37$, and ATSPPH-V, $r = .23$, both $p < .001$)[54]. The eHEALS was developed to assess an individual’s perceived knowledge, confidence, and skills in finding, evaluating, and applying electronic health information[35]. However, it is not a direct measure of attitudes toward AI. The ATSPPH subscales represent more conceptually distant constructs, as they assess attitudes toward seeking professional psychological help when experiencing psychological or emotional difficulties[55]. The low HTMT values and the finding that AVE exceeded MSV suggest that AIAS-4 scores were empirically distinguishable from these other constructs rather than measuring the same underlying concept[46, 47]. Taken together, these findings provide preliminary validity evidence

based on relations with external variables and help position AIAS-4 within a broader nomological network. They should not be interpreted as evidence that eHEALS or ATSPPH measures the same construct as attitudes toward AI.

Measurement invariance evidence differed by grouping variable. Across age groups, fit did not deteriorate at any level and $\Delta\chi^2$ tests were nonsignificant, supporting strict invariance. Across gender, the configural model fit well and *ΔCFI* at the metric step was within conventional bounds (− .004), but *RMSEA* and *SRMR* increased, and the improvement at the scalar and strict steps does not fully resolve this ambiguity. Because fit indices differ in their sensitivity to noninvariance, no single index should be decisive[48, 49]; metric invariance across gender should therefore be regarded as tentative pending replication with larger male subsamples. Invariance across demographic groups within Vietnam is also conceptually distinct from cross-linguistic equivalence, which requires multi-country data such as those of Talik et al.[56] for Poland, the United Kingdom, and the United States. Other adaptations have likewise reported supportive but not uniform invariance findings[30, 57].

Relative test–retest reliability was good (ICC = .84), indicating stable rank ordering over two weeks[42], and the mean change of .12 points (*d* = .12) showed no systematic shift. Measurement error, however, is a distinct property: the SEM of .68 and MDC_{95} of 1.90 points indicate that an individual change smaller than about two scale points cannot be distinguished from error, a statistical rather than a clinical threshold[58]. The Bland–Altman limits (− 1.78 to 2.02) revealed individual variation not captured by the ICC[59], and the negative proportional bias (*b* = − .172, *p* < .0) means that disagreement between occasions depended on score level. Whereas previous adaptations have typically reported a single stability coefficient[30], jointly reporting relative reliability, measurement error, and absolute agreement gives a fuller account of score reproducibility and cautions against interpreting small individual changes.

The four-item format reduces respondent burden and suits large surveys and multivariable models of technology acceptance and AI-related learning[18, 60]. Higher scores indicate more favorable general evaluations of AI; they do not imply low AI anxiety, few perceived risks, greater AI literacy, higher trust, or comprehensive technology acceptance[13–15, 17]. Although brevity also makes the scale attractive for intervention studies, responsiveness was not evaluated here, so the Vietnamese AIAS-4 cannot yet be recommended for detecting intervention effects[21, 61].

Several limitations apply. First, participants were recruited online through nonprobability sampling, which may introduce selection bias and limit generalizability[62]; the sample was also predominantly female and uneven across fields of study, so future studies should recruit more balanced samples across gender, discipline, and region. Second, the linguistic review was conducted with experts rather than through cognitive interviews with students, so respondent-level interpretation of the items, particularly the broad wording of Items 1 and 4, was not directly examined; future adaptations should include cognitive interviews with the target population[33]. Third, no comparator measure of AI attitudes was included, limiting convergent validity evidence; the AIAS-4 should next be examined alongside other AI attitude scales[11, 63], AI anxiety[64], trust in AI[65], and AI literacy, as proposed in the original study[18]. Fourth, invariance was tested only across demographic groups within one country; cross-linguistic equivalence remains to be established through multigroup CFA or differential item functioning analyses across language groups. Fifth, the retest subsample was small (*n* = 60) and showed proportional bias, so individual change scores should be interpreted cautiously until replicated in larger samples spanning the full score range. Finally, responsiveness was not assessed; longitudinal studies with external criteria of change are needed before the scale is used to evaluate interventions[58].

Conclusion

The Vietnamese AIAS-4 retained the one-factor structure of the original scale across two independent subsamples, showed high internal consistency and good two-week test–retest reliability, and was invariant across age groups and, less consistently at the metric level, across gender. Evidence from relations with external variables was modest, and proportional bias in the retest analysis cautions against interpreting small individual changes. The scale is a promising

brief measure of general positive attitudes toward AI among Vietnamese university students; its responsiveness and cross-linguistic equivalence remain to be established.

Declarations

Competing interests

The authors declare no competing interests.

Supplementary Information

Supplementary Information is available for this paper.

Funding

This research received no specific grant from any funding agency.

Author Contribution

All authors approved the final manuscript for submission. Thi Duyen Vu contributed to the conceptualization and coordination of the study, methodology, data collection, project administration, and preparation and revision of the manuscript. Tuan Kiet Pham contributed to data analysis and visualization, translation and cultural adaptation, and preparation of the manuscript. Thi Bao Ngan Nguyen contributed to the preparation of the manuscript, and review and editing, and research support. An Dinh Duy Phan contributed to data analysis and preparation of the manuscript. Minh Thuan Ly and Huu Son Hai Le contributed to the preparation of the manuscript. Simone Grassini contributed to methodological supervision, validation, and revision of the manuscript and served as the corresponding author. All authors contributed to the development of the theoretical framework, critically reviewed the manuscript, and approved the final version for submission.

Acknowledgement

The authors would like to extend our sincere thanks to all study participants and the eight experts who reviewed the translation included in this study.

Data Availability

The datasets generated and analysed during the current study are not publicly available because participants did not consent to public data sharing but are available from the corresponding author on reasonable request.

References

1. Acosta-Enriquez, B. G., Arbulu Ballesteros, M. A., Jordan, H. & Lopez Roca, O. Saavedra Tirado, K. Analysis of college students' attitudes toward the use of ChatGPT in their academic activities: effect of intent to use, verification of information and responsible use. *BMC Psychol.* **12**, 255. 10.1186/s40359-024-01764-z (2024).

2. Wu, F., Dang, Y. & Li, M. A. Systematic Review of Responses, Attitudes, and Utilization Behaviors on Generative AI for Teaching and Learning in Higher Education. *Behav. Sci. (Basel)*. **15**, 467. 10.3390/bs15040467 (2025).
3. Chan, C. K., Hu, W. & Y. & Students' voices on generative AI: perceptions, benefits, and challenges in higher education. *Int. J. Educational Technol. High. Educ.* **20**, 43. 10.1186/s41239-023-00411-8 (2023).
4. Sun, Y. et al. AI literacy and psychosocial factors shaping Chinese university students' attitudes and behavioral intentions toward generative AI use. *BMC Psychol.* **14**, 236. 10.1186/s40359-026-03989-6 (2026).
5. Shahzad, M. F., Xu, S. & Asif, M. Factors affecting generative artificial intelligence, such as ChatGPT, use in higher education: An application of technology acceptance model. *Br. Edu. Res. J.* **51**, 489–513. 10.1002/berj.4084 (2024).
6. Teo, T., Scherer, R., Fung, A. S. K. & Fung, C. S. L. Factors associated with students' adoption of artificial intelligence technology in tertiary education: A meta-analytic review. *Educational Res. Rev.* **52** 10.1016/j.edurev.2026.100804 (2026).
7. Ibrahim, F., Munscher, J. C., Daseking, M. & Telle, N. T. The technology acceptance model and adopter type analysis in the context of artificial intelligence. *Front. Artif. Intell.* **7**, 1496518. 10.3389/frai.2024.1496518 (2024).
8. Santos-Jaén, J. M., León-Gómez, A. & Valls Martínez, M. C. Llorente Muñoz, V. University students' perceptions of the impact of artificial intelligence in the business sector on their educational and professional development. *Educ. Inform. Technol.* **30**, 24395–24428. 10.1007/s10639-025-13712-4 (2025).
9. Eagly, A. H. & Chaiken, S. *The psychology of attitudes* (Harcourt Brace Jovanovich College, 1993).
10. Kim, S. W. & Lee, Y. Investigation into the Influence of Socio-Cultural Factors on Attitudes toward Artificial Intelligence. *Educ. Inform. Technol.* **29**, 9907–9935. 10.1007/s10639-023-12172-y (2023).
11. Schepman, A. & Rodway, P. Initial validation of the general attitudes towards Artificial Intelligence Scale. *Comput. Hum. Behav. Rep.* **1**, 100014. 10.1016/j.chbr.2020.100014 (2020).
12. Novotny, M., Weber, W., Kern, C. & Kreuter, F. Measuring public opinion towards artificial intelligence: development and validation of a general AI attitude short scale. *AI Soc.* **41**, 4227–4260. 10.1007/s00146-025-02478-5 (2025).
13. Kaya, F. et al. The Roles of Personality Traits, AI Anxiety, and Demographic Factors in Attitudes toward Artificial Intelligence. *Int. J. Human–Computer Interact.* **40**, 497–514. 10.1080/10447318.2022.2151730 (2022).
14. Kelly, S., Kaye, S. A. & Oviedo-Trespalacios, O. What factors contribute to the acceptance of artificial intelligence? A systematic review. *Telematics Inform.* **77** 10.1016/j.tele.2022.101925 (2023).
15. Ng, D. T. K., Leung, J. K. L., Chu, S. K. W. & Qiao, M. S. Conceptualizing AI literacy: An exploratory review. *Computers Education: Artif. Intell.* **2** 10.1016/j.caeai.2021.100041 (2021).
16. Davis, F. D. & Perceived Usefulness Perceived Ease of Use, and User Acceptance of Information Technology. *Manage. Inform. Syst. Q.* **13**, 319–340. 10.2307/249008 (1989).
17. Koenig, P. D. Attitudes toward artificial intelligence: combining three theoretical perspectives on technology acceptance. *AI Soc.* **40**, 1333–1345. 10.1007/s00146-024-01987-z (2024).
18. Grassini, S. Development and validation of the AI attitude scale (AIAS-4): a brief measure of general attitude toward artificial intelligence. *Front. Psychol.* **14**, 1191628. 10.3389/fpsyg.2023.1191628 (2023).
19. DeVellis, R. F. *Scale development: Theory and applications* Vol. 26 (SAGE Publications, Inc, 2012).
20. Epstein, J., Santo, R. M. & Guillemin, F. A review of guidelines for cross-cultural adaptation of questionnaires could not bring out a consensus. *J. Clin. Epidemiol.* **68**, 435–441. 10.1016/j.jclinepi.2014.11.021 (2015).
21. Prinsen, C. A. C. et al. COSMIN guideline for systematic reviews of patient-reported outcome measures. *Qual. Life Res.* **27**, 1147–1157. 10.1007/s11136-018-1798-3 (2018).
22. Vandenberg, R. J. & Lance, C. E. A Review and Synthesis of the Measurement Invariance Literature: Suggestions, Practices, and Recommendations for Organizational Research. *Organizational Res. Methods*. **3**, 4–70. 10.1177/109442810031002 (2000).

23. Si, J. & Exploring, A. I. literacy, attitudes toward AI, and intentions to use AI in clinical contexts among healthcare students in Korea: a cross-sectional study. *BMC Med. Educ.* **25**, 1233. 10.1186/s12909-025-07766-8 (2025).
24. Naiseh, M. et al. The Interplay of Self-Efficacy, Well-Being, and Competency. *J. Technol. Behav. Sci.* **10**, 705–718. 10.1007/s41347-025-00486-2 (2025). Attitudes Towards AI.
25. Sindermann, C. et al. Assessing the Attitude Towards Artificial Intelligence: Introduction of a Short Measure in German, Chinese, and English Language. *KI - Künstliche Intelligenz*. **35**, 109–118. 10.1007/s13218-020-00689-0 (2020).
26. Stein, J. P., Messingschlager, T., Gnambs, T., Hutmacher, F. & Appel, M. Attitudes towards AI: measurement and associations with personality. *Sci. Rep.* **14**, 2909. 10.1038/s41598-024-53335-2 (2024).
27. Köse, N., Şimşek, E. & Demir, M. C. Adaptation of Artificial Intelligence Attitude Scale (AIAS-4) into Turkish: a validity and reliability study. *Curr. Psychol.* **44**, 8096–8105. 10.1007/s12144-025-07418-6 (2025).
28. Luque-Reca, O., Peñacoba, C., Catalá, P., Gutiérrez Hermoso, L. & Grassini, S. Artificial intelligence attitudes in higher education: Spanish validation of the AIAS-4 and its associations with well-being and academic AI usage. *Curr. Psychol.* **45** 10.1007/s12144-025-08566-5 (2026).
29. Møgelvang, A. & Grassini, S. Validating the AI attitude scale (AIAS-4) and exploring attitudinal differences in a large sample of Norwegian university students. *Discover Educ.* **4** 10.1007/s44217-025-00657-6 (2025).
30. Satici, S. A., Okur, S., Yilmaz, F. B. & Grassini, S. Psychometric properties and Turkish adaptation of the artificial intelligence attitude scale (AIAS-4): evidence for construct validity. *BMC Psychol.* **13**, 297. 10.1186/s40359-025-02505-6 (2025).
31. Truong, N. M., Vo, T. Q., Tran, H. T. B., Nguyen, H. T. & Pham, V. N. H. Healthcare students' knowledge, attitudes, and perspectives toward artificial intelligence in the southern Vietnam. *Heliyon* **9**, e22653. 10.1016/j.heliyon.2023.e22653 (2023).
32. Tien, L. M. Exploring undergraduate students' general attitudes towards Artificial Intelligence: A perspective from Vietnam. *J. Lang. Cult. Educ.* **12**, 16–22. 10.2478/jolace-2024-0014 (2024).
33. Cruchinho, P. et al. Cross-Cultural Adaptation, and Validation of Measurement Instruments: A Practical Guideline for Novice Researchers. *J. Multidiscip Healthc.* **17**, 2701–2728. 10.2147/JMDH.S419714 (2024). Translation.
34. Polit, D. F., Beck, C. T. & Owen, S. V. Is the CVI an acceptable indicator of content validity? Appraisal and recommendations. *Res. Nurs. Health*. **30**, 459–467. 10.1002/nur.20199 (2007).
35. Norman, C. D., Skinner, H. A. & eHEALS The eHealth Literacy Scale. *J. Med. Internet Res.* **8**, e27. 10.2196/jmir.8.4.e27 (2006).
36. Fischer, E. H. & Farina, A. Attitudes toward seeking professional psychological help: a shortened form and considerations for research. *J. Coll. student Dev.* **36**, 368–373 (1995).
37. Lorenzo-Seva, U. SOLOMON: a method for splitting a sample into equivalent subsamples in factor analysis. *Behav. Res. Methods*. **54**, 2665–2677. 10.3758/s13428-021-01750-y (2022).
38. Hayton, J. C., Allen, D. G. & Scarpello, V. Factor Retention Decisions in Exploratory Factor Analysis: a Tutorial on Parallel Analysis. *Organizational Res. Methods*. **7**, 191–205. 10.1177/1094428104263675 (2004).
39. Watkins, M. W. Exploratory Factor Analysis: A Guide to Best Practice. *J. Black Psychol.* **44**, 219–246. 10.1177/0095798418771807 (2018).
40. Rosseel, Y. lavaan: AnRPackage for Structural Equation Modeling. *J. Stat. Softw.* **48** 10.18637/jss.v048.i02 (2012).
41. Hair, J. F., Black, W. C., Babin, B. J. & Anderson, R. E. *Multivariate data analysis* (Cengage Learning EMEA, 2019).
42. Koo, T. K. & Li, M. Y. A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. *J. Chiropr. Med.* **15**, 155–163. 10.1016/j.jcm.2016.02.012 (2016).

43. Bland, J. M. & Altman, D. G. Measuring agreement in method comparison studies. *Stat. Methods Med. Res.* **8**, 135–160. 10.1177/096228029900800204 (1999).
44. Mokkink, L. B. et al. COSMIN Risk of Bias tool to assess the quality of studies on reliability or measurement error of outcome measurement instruments: a Delphi study. *BMC Med. Res. Methodol.* **20**, 293. 10.1186/s12874-020-01179-5 (2020).
45. Cristobal, E., Flavián, C. & Guinalíu, M. Perceived e-service quality (PeSQ). *Managing Service Quality: Int. J.* **17**, 317–340. 10.1108/09604520710744326 (2007).
46. Fornell, C. & Larcker, D. F. Evaluating Structural Equation Models with Unobservable Variables and Measurement Error. *J. Mark. Res.* **18** 10.2307/3151312 (1981).
47. Henseler, J., Ringle, C. M. & Sarstedt, M. A new criterion for assessing discriminant validity in variance-based structural equation modeling. *J. Acad. Mark. Sci.* **43**, 115–135. 10.1007/s11747-014-0403-8 (2014).
48. Chen, F. F. Sensitivity of Goodness of Fit Indexes to Lack of Measurement Invariance. *Struct. Equation Modeling: Multidisciplinary J.* **14**, 464–504. 10.1080/10705510701301834 (2007).
49. Putnick, D. L. & Bornstein, M. H. Measurement Invariance Conventions and Reporting: The State of the Art and Future Directions for Psychological Research. *Dev. Rev.* **41**, 71–90. 10.1016/j.dr.2016.06.004 (2016).
50. ITC. ITC Guidelines for Translating and Adapting Tests. (Second Edition). *International Journal of Testing* 18, 101–134, (2017). 10.1080/15305058.2017.1398166
51. Martinent, G., Guillet-Descas, E. & Moiret, S. Reliability and validity evidence for the French Psychological Need Thwarting Scale (PNTS) scores: Significance of a distinction between thwarting and satisfaction of basic psychological needs. *Psychol. Sport Exerc.* **20**, 29–39. 10.1016/j.psychsport.2015.04.005 (2015).
52. Knekta, E., Runyon, C. & Eddy, S. One Size Doesn't Fit All: Using Factor Analysis to Gather Validity Evidence When Using Surveys in Your Research. *CBE Life Sci. Educ.* **18**, rm1. 10.1187/cbe.18-04-0064 (2019).
53. McNeish, D. Thanks coefficient alpha, we'll take it from here. *Psychol. Methods*. **23**, 412–433. 10.1037/met0000144 (2018).
54. Gignac, G. E. & Szodorai, E. T. Effect size guidelines for individual differences researchers. *Pers. Indiv. Differ.* **102**, 74–78. 10.1016/j.paid.2016.06.069 (2016).
55. Fischer, E. H. & Turner, J. L. Orientations to seeking professional help: development and research utility of an attitude scale. *J. Consult Clin. Psychol.* **35**, 79–90. 10.1037/h0029636 (1970).
56. Talik, W., Talik, E. B. & Grassini, S. Measurement invariance of the artificial intelligence attitude scale (AIAS-4): cross-cultural studies in Poland, the USA, and the UK. *Curr. Psychol.* **44**, 15758–15766. 10.1007/s12144-025-08348-z (2025).
57. Morales-Garcia, W. C., Sairitupa-Sanchez, L. Z., Morales-Garcia, S. B. & Morales-Garcia, M. Adaptation and Psychometric Properties of an Attitude toward Artificial Intelligence Scale (AIAS-4) among Peruvian Nurses. *Behav. Sci. (Basel)*. 14. 10.3390/bs14060437 (2024).
58. Mokkink, L. B. et al. The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. *J. Clin. Epidemiol.* **63**, 737–745. 10.1016/j.jclinepi.2010.02.006 (2010).
59. Bland, J. M. & Altman, D. G. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* **1**, 307–310. 10.1016/S0140-6736(86)90837-8 (1986).
60. Lantos, D. et al. The performance of long vs. short questionnaire-based measures of depression, anxiety, and psychological distress among UK adults: A comparison of the patient health questionnaires, generalized anxiety disorder scales, malaise inventory, and Kessler scales. *J. Affect. Disord*. **338**, 433–439. 10.1016/j.jad.2023.06.033 (2023).

61. Mokkink, L., Terwee, C. & de Vet, H. Key concepts in clinical epidemiology: Responsiveness, the longitudinal aspect of validity. *J. Clin. Epidemiol.* **140**, 159–162. 10.1016/j.jclinepi.2021.06.002 (2021).
62. Oudat, Q. & Bakas, T. Merits and Pitfalls of Social Media as a Platform for Recruitment of Study Participants. *J. Med. Internet Res.* **25**, e47705. 10.2196/47705 (2023).
63. Schepman, A. & Rodway, P. The General Attitudes towards Artificial Intelligence Scale (GAAIS): Confirmatory Validation and Associations with Personality, Corporate Distrust, and General Trust. *Int. J. Human–Computer Interact.* **39**, 2724–2741. 10.1080/10447318.2022.2085400 (2022).
64. Wang, Y. Y. & Wang, Y. S. Development and validation of an artificial intelligence anxiety scale: an initial application in predicting motivated learning behavior. *Interact. Learn. Environ.* **30**, 619–634. 10.1080/10494820.2019.1674887 (2019).
65. McGrath, M. J., Lack, O., Tisch, J. & Duenser, A. Measuring trust in artificial intelligence: validation of an established scale and its short form. *Front. Artif. Intell.* **8**, 1582880. 10.3389/frai.2025.1582880 (2025).

Figures

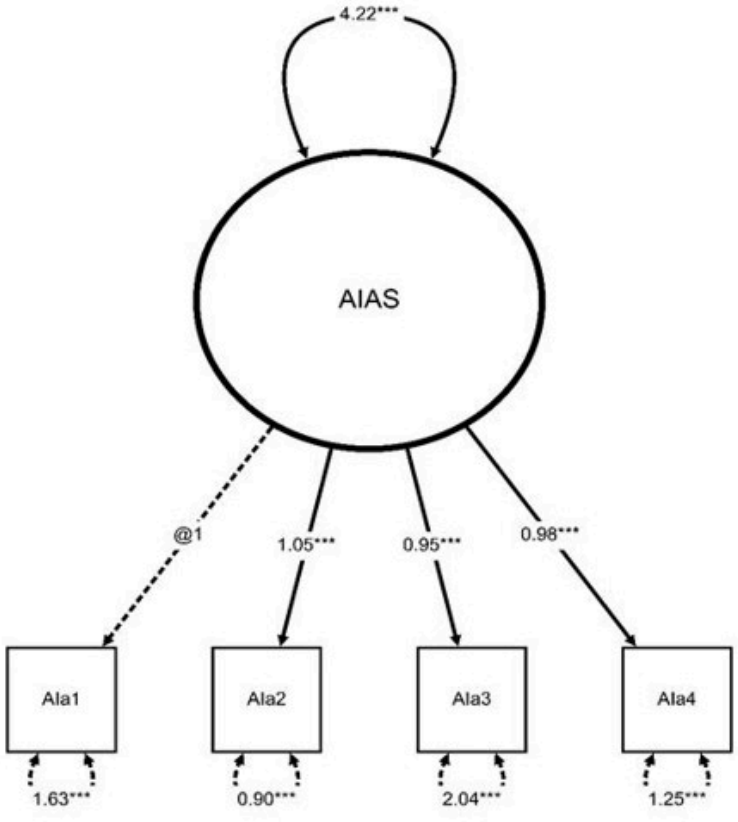


Figure 1

Path diagram of the one-factor confirmatory model of the Vietnamese AIAS-4 (n = 349). Values on arrows are standardized factor loadings; values on the right are residual variances. All loadings were significant at $p < .001$.

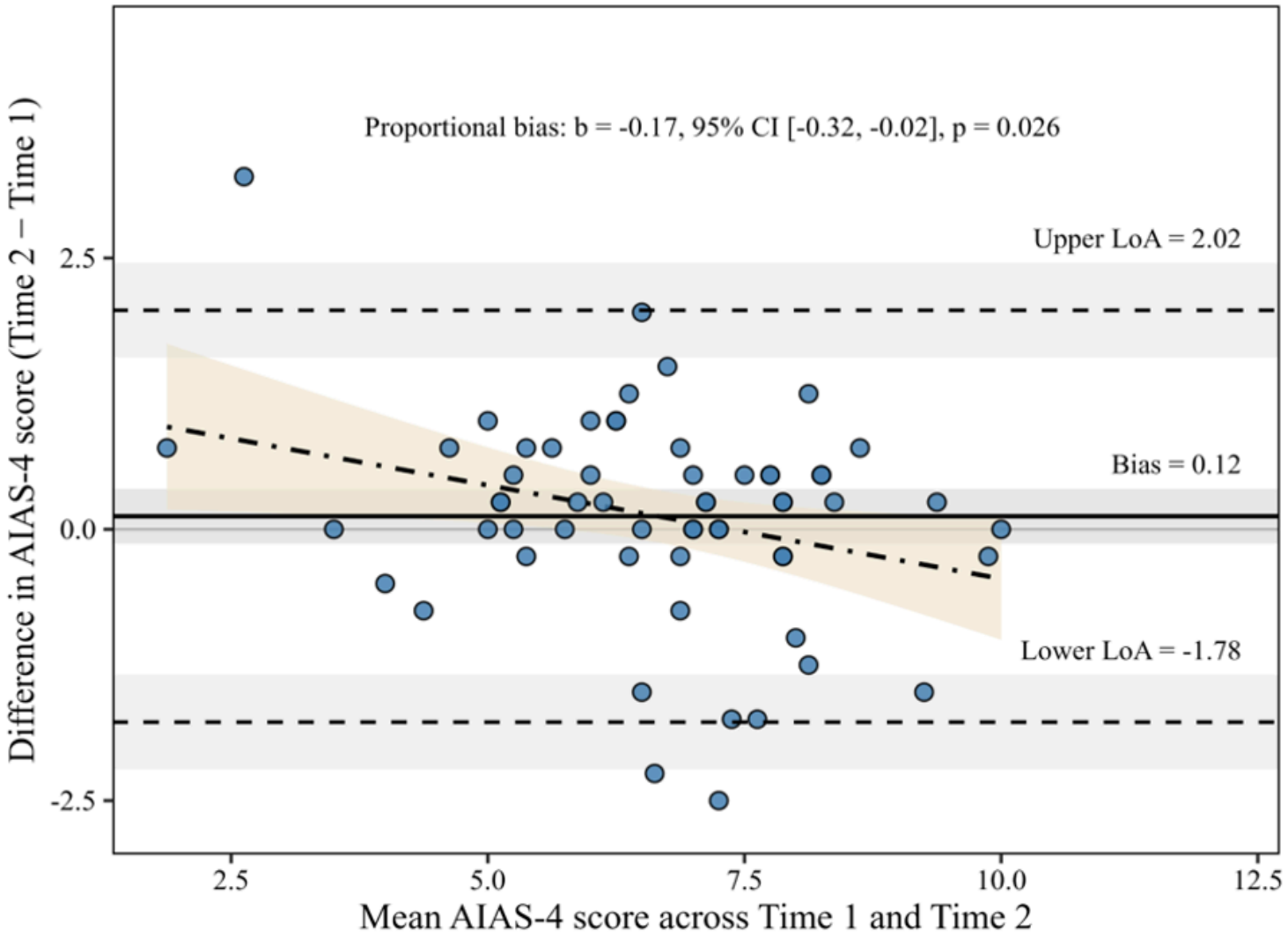


Figure 2

Bland–Altman plot of test–retest agreement for the Vietnamese AIAS-4 (*n*= 60). Differences between Time 2 and Time 1 scores are plotted against their mean. The solid line indicates the mean bias (.12), dashed lines the 95% limits of agreement (−1.78 and 2.02), and the sloped line the regression of differences on means, showing proportional bias (*b* = −.172, *p* = .026).

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- AIAS4VietnameseSupplementaryInformation04102026.docx