Current Psychology (2026) 45:470
https://doi.org/10.1007/s12144-025-08736-5

Reliability generalization of the artificial intelligence anxiety scale

Yıldız Yıldırım[3] · Tuba Gündüz[1] · Melek Gülşah Şahin[2]

Received: 29 July 2024 / Accepted: 8 October 2025
© The Author(s) 2026

Abstract

With the rapid development of artificial intelligence (AI), AI anxiety is becoming a psychological construct that is associated with a number of other variables and is frequently reported in a variety of populations. This study aims to examine the reliability of the original version by Wang and Wang and the adapted version by Akkaya et al. of Artificial Intelligence Anxiety Scale, a prominent measure of AI anxiety. Meta-analytic reliability generalization of the original version of the AIAS was performed using Cronbach's alpha coefficients from 15 primary studies with a total sample size of 4975. For the the adapted AIAS version by Akkaya et al., the total sample size was 3407 and the number of primary studies was 10. Categorical moderator analysis of the obtained overall Cronbach's alpha coefficients according to Likert type, research type, scale language, study group, country, and study field variables was performed. Also, continuous moderator analysis was conducted for the ratio of women variable. When the overall alpha coefficients were examined, it was seen that all pooled Cronbach's alpha coefficients were high. The pooled alpha coefficient for the overall scale of the original version of the AIAS is 0.949 [%95 CI = 0.944—0.963], while this value is 0.878 [%95 CI = 0.844—0.935] for the adapted AIAS version by Akkaya et al. Even the pooled alpha coefficients obtained in different variable categories were generally high. For the original 21-item AIAS, when the moderator analysis results are examined, it is concluded that the difference between the overall alpha coefficients in the country and study field variables is significant. In addition, it was determined that the ratio of women variable significantly predicted the alpha coefficient for the Job Replacement subscale. Finally, suggestions for future research and practice are presented along with relevant implications.

Introduction

The rapid development of technology in recent years has led to the profound transformations in human history. The most recent of these developments is in the field of Artificial Intelligence (AI), which combines machine and human characteristics. The term AI was first used by John McCarthy in 1955 to describe the ability of a machine to use human characteristics such as perception, reasoning, learning, or communication (Chan & Zary, 2019). According to McCarthy (2007), English mathematician Alan Turing was the first person to work in the field of AI and stated that AI could be researched by programming computers. In defining the concept of AI, science and engineering played a role together in making intelligent machines. Today, AI appears in many fields such as health, defense, industry, education, sports, tourism, accounting, and improvements are being made in the processes day by day.

In the face of developments in the field of AI, individuals can access a lot of information, process information easily, improve their quality of life, reduce human errors, and use their time efficiently. Every day, individuals are discovering or learning from others about the conveniences that AI provides them. This allows AI to develop furtheror AI-human interaction continues to increase. AI is not only utilized by individuals, but companies are also using AI to increase the efficiency of their processes and design new products (Colosimo et al., 2021). Although there are positive developments related to AI, there are a number of ethical, employment,

✉ Tuba Gündüz
tubagunduz@mu.edu.tr

Yıldız Yıldırım
yildizyldrm@gmail.com

Melek Gülşah Şahin
mgulsahsahin@gazi.edu.tr

[1] Faculty of Education, Muğla Sıtkı Koçman University, Muğla, Türkiye

[2] Gazi Faculty of Education, Gazi University, Ankara, Türkiye

[3] Faculty of Education, Aydın Adnan Menderes University, Aydın, Türkiye

Published online: 17 February 2026

sociopolitical, and economic concerns, as well as concerns regarding human reasoning and thinking reduction, fear of human domination, fear of being used for malicious purposes, security, deterioration in psychological and physical health, and prejudices about AI. (Comer, 2023; Gull et al., 2023; Kim et al., 2023; Neudert et al., 2020; Pektaş, 2024; Tugay & Tugay, 2019; Rayhan & Rayhan, 2023; Wang & Wang, 2019). In short, the rapid inclusion of AI in our lives does not only have a positive contribution.

Due to the long history of AI, machine-semi-machine-human interactions have also been encountered in science fiction movies and novels. There are even fears and anxiety about AI that evoke these scenarios in individuals (Comer, 2023; Wang et al., 2022). The concept of AI anxiety appears. AI-related anxiety is an emerging and not yet fully defined phenomenon that may develop through individuals' interactions with artificial intelligence. Comparable forms of this anxiety include computer-related and robot-induced anxiety. (Wang & Wang, 2022) And also the term AI anxiety refers to the fear and trepidation being expressed about AI becoming uncontrollable (Johnson & Verdicchio, 2017). Wang et al. (2024), in their study, used the concept of AI anxiety as "GenAI (Generative artificial intelligence) anxiety" and defined it as a comprehensive emotional state that includes fear of using GenAI and concerns about its potential threats. In general, AI anxiety can be defined under 3 headings: Special focus on AI programs to the exclusion of human impact, failure to develop an understanding of human and program autonomy, and misunderstanding of technological developments. Accordingly, the majority of concerns about AI are based on misunderstanding and confusion (Johnson & Verdicchio, 2019). Considering the current interest of AI in every field, it is also important to examine the psychological effects of artificial intelligence on individuals (Kim et al., 2023). Thus, it will be possible to develop strategies to ensure a positive mood in individuals (Rayhan & Rayhan, 2023). Positive attitudes towards artificial intelligence in individuals can reduce anxiety towards information technology and increase its use (Kwak et al., 2022).

Studies on AI anxiety and measurement tools to assess the level of AI anxiety are available in the literature. The first of these is the AI Anxiety Scale (AIAS). AI Anxiety Scale (AIAS) developed by Wang and Wang (2022) for AI anxiety, which is an important issue today. AIAS consists of 21 items and four subscales. The scale items that prepared to assess individuals' anxiety about AI-related products were scored using a 7-point (1 = never through 7 = completely). The subscales are Learning (8 items), AI Configuration (3 items), Job Replacement (6 items), Sociotechnical Blindness (4 items). The Learning subscale includes items that related with learning about AI techniques or products with anxiety. The Job Replacement subscale includes items related to the negative impact of the widespread use of AI techniques or products on people's professional responsibility and concern about losing their jobs. Sociotechnical Blindless subscale includes items that express concern that the use of AI techniques or products may cause situations that cannot be controlled by humans. AI Configuration subscale includes items on whether individuals find humanoid AI techniques or products scary or intimidating. The original validity and reliability studies of the AIAS were conducted with 301 participants. The majority of the participants were men (65.1%), between the ages of 21–30 (48.8%), had a Bachelor's degree (60.2%) and were students (39.2%). Cronbach's alpha reliability values were obtained to estimate the reliability of the overall scale and subscales. The overall reliability value of the scale was 0.964. The reliability of each of the subscales were as follows: Learning = 0.974; Job Replacement = 0.917; Sociotechnical Blindness = 0.917; and AI Configuration = 0.961. Criterion validity was calculated and a value of 0.864 was obtained. The minimum within-item correlations were Learning = 0.624; Job Replacement = 0.480; Sociotechnical Blindness = 0.703; and AI Configuration = 0.860. Inter-factor correlations ranged between 0.494 and 0.725. Akkaya et al. (2021) adapted the AIAS developed by Wang and Wang (2022) into Turkish. However, in the adaptation process, items 6, 7, 8, 9, and 10, which had factor loadings below 0.30 or tended to load on more than one factor after EFA, were removed from the scale one by one and the scale was reduced from 21 to 16 items. Although the subscale names of the scale did not change, the number of items in the Learning and Job Replacement subscales decreased by three and two, respectively, compared to original form. In the Turkish form, in the Job Replacement subscale, two items expressing individuals' concern that AI techniques or products may make people dependent and lazy were removed. In the Learning dimension, statements that included concerns about taking courses, reading guides and following developments in AI techniques or products were removed. Cronbach's alpha reliability values were obtained to estimate the reliability of the scale. The reliability of each of the four subscales was: Learning = 0.948; Job Replacement = 0.895; Sociotechnical Blindness = 0.875; and AI Configuration = 0.950. A coefficient alpha of 0.937 was calculated for total scale.

Another measurement tool was developed by Li and Huang (2020) to assess the level of AI anxiety. Li and Huang (2020), defined eight factors to explain the source of AI anxiety by using integrated fear acquisition theory with 23 items that were scored on a Likert scale from 1 (strongly disagree) to 7 (strongly agree): (1) Privacy Violation Anxiety, (2) Bias Behavior Anxiety, (3) Job Replacement

Anxiety, (4) Learning Anxiety, (5) Existential Risk Anxiety, (6) Ethics Violation Anxiety, (7) Artificial Consciousness Anxiety and (8) Lack of Transparency Anxiety. And also, in their model there were four factors at the second-order level that combine the factors in pairs, respectively: Conditioning (1–2), Vicarious Exposure (3–4), Transmission of Information and Instruction (5–6), Inability to Recall a Pertinent Experience (7–8). Validity and reliability studies were conducted with 492 participants. The majority of the participants were between the ages of 18–22 (55.2%) and men (53.8%). No other demographic information about the participants was reported.

In this study, it is aimed to conduct the meta-analytic reliability generalization (RG) of the AIAS (Wang & Wang, 2022). The fact that AIAS is an older scale as of the date of its online publication (2019) was also considered important in its selection. It is seen that the use of AIAS in different studies started in 2020 (Dai et al., 2020). There are two different Turkish adaptation studies of the scale (Akkaya et al., 2021; Terzi, 2020). In addition, when the demographic information about the participants of the validity and reliability study of AIAS was examined, it was determined that it was carried out in different educational levels and different occupation groups (manufacturing, service, student, government, education/research, medical, other). Considering that the widespread use of AI will diversify the users, it is also important to collect data from individuals with different levels of education and individuals in different occupational groups. Also based on the databases analyzed in the study; it was concluded that AIAS is the most widely used assessment tool for artificial intelligence anxiety. Within the scope of the study, general reliability for the overall scale and subscales will be estimated for both versions of the AIAS scale. In addition, how the reliability values change within the scope of the selected categorical and continuous variables will also be examined. In determining the variables considered within the scope of the study, the variables affecting the reliability value and examined in the literature in the generalization of meta-analytic reliability were taken into consideration. Accordingly, likert type (Şahin & Yıldırım, 2024; Şen, 2021), scale language (Orhan, 2022; Rubio-Aparicio et al., 2020; Şahin & Yıldırım, 2024; Yörük & Şen, 2023), country (Blázquez-Rincón et al., 2022; López-Pina et al., 2015; Orhan, 2022; Şen, 2021), study group (Eser & Doğan, 2023; Özdemir et al., 2020; Şahin & Yıldırım, 2024; Yörük & Şen, 2023) research type (Batı & Irmak, 2024; López-Pina et al., 2015; Şahin & Yıldırım, 2024) and study field (Özdemir et al., 2020; Rubio-Aparicio et al., 2020; Şahin & Yıldırım, 2024) variables were considered as categorical variables. The ratio of women (Beretvas et al., 2002; Lenz et al., 2024; Şahin & Yıldırım, 2024; Yörük & Şen, 2023) was determined as a continuous variable.

Methodology

In this study, meta-analytic reliability generalization of Cronbach's alpha reliability coefficients of AIAS was conducted. Accordingly, both the 21-itemoriginal AIAS published online in 2019 (Wang & Wang, 2022) and the 16-item Turkish version (Akkaya et al., 2021) were included in the analyses as the overall scale. Cronbach's alpha reliability coefficients for the overall scale and subscales of Learning, Job Replacement, Sociotechnical Blindness and AI Configuration for the original AIAS by Wang and Wang were analyzed by Vacha-Haase's (2008) meta-analytic reliability generalization. Reliability generalization was also performed for the overall scale, subscales of Learning and Job Replacement of the adapted AIAS version by Akkaya et al. Since the items in the Sociotechnical Blindness and AI Configuration subscales were the same in this adapted version and the original version, separate analyses were not conducted for these subscales. However, scale version (21-item and 16-item) was considered as a categorical moderator variable for these subscales.

The REGEMA Checklist was followed in this study. The REGEMA checklist consists of the features that should be reported in the title, abstract, introduction, method, results, discussion, funding and protocol of a meta-analytic reliability generalization study. This checklist consists of 30 items and the responses are categorized as "Yes", "No", "Unclear" and "NA" (Sánchez-Meca et al., 2021). Unlike other reporting guidelines, the REGEMA checklist directly guides meta-analytic reliability generalization studies. The REGEMA checklist was used to ensure the correct construction and reporting this meta-analytic reliability generalization. Thus, we planned to provide more effective results about the reliability coefficients of AIAS for future studies and researchers who will apply the scale.

Study selection criteria

This study aimed to estimate pooled Cronbach's alpha of the overall scale and subscales of the original AIAS by Wang and Wang and the adapted AIAS version by Akkaya et al. via meta-analytic reliability generalization. The criteria for inclusion in the meta-analysis are as follows:

- Studies should be included regardless of sample size, target population, publication year, publication language and publication type.
- The primary studies must have administered the original AIAS by Wang and Wang and its subscales or the adapted AIAS version by Akkaya et al. and its subscales and reported Cronbach's alpha coefficient.
- The sample size must be reported in the primary study.

In addition to these criteria, in order to prevent language bias, it was examined whether the scanned studies used the AIAS through translation programs regardless of language. However, no study was found to use any version of the scale in languages other than English and Turkish. In order to reduce publication bias, primary studies were not limited to articles. Other sources such as theses, books reports, reports were also examined according to the criteria. Thus, it was aimed to deal with gray literature that includes sources from research between the file drawer and publication, or from obscure publishing organizations (Yıldırım, 2023).

Search strategies and data extraction

Web of Science, Google Scholar, SCOPUS, ERIC, Taylor & Francis, ProQuest databases, which are considered important in the field of education and psychology, were screened with the keyword "Artificial Intelligence Anxiety Scale". All fields of the databases were screened, not only the title and abstract fields. As a result of the screening, as of May 27, 2024, a total of 245 primary studies were examined. Primary studies that used the scale but could not be reached were requested via e-mail by the corresponding authors. For the primary studies that were accessible but did not report Cronbach's α, we sent an e-mail to the corresponding author or made a phone call to those whose phone numbers we could access. All of these studies were examined through their full texts and 26 studies were included in the meta-analytic reliability generalization.

The REGEMA Checklist was followed in this study. The REGEMA flowchart on the process of selection of primary studies according to the inclusion criteria is presented in Fig. 1.

In Fig. 1, these 26 studies reported Cronbach's alpha (α) coefficient for at least one of the original AIAS by Wang and Wang and its subscales or the adapted AIAS version by Akkaya et al. and its subscales. According to the REGEMA flowchart given in Fig. 1, the initial search yielded 245 studies. Of these 245 studies, 83 duplicate were eliminated. After the removal of duplications, 10 of the studies were eliminated because they were theoretical and 4 of the studies were systematic reviews. The full text of 20 of the studies could not be reached and the authors were tried to be reached via e-mail but no or late feedback was received. No scale was used in 78 of the 128 studies whose full text was accessed, and a scale different from the AIAS scale was used in 12 of them. Since reliability values were not reported in 12 of the remaining 38 studies, these studies were eliminated. Following the elimination of these studies, 26 studies were included in this reliability generalization meta-analysis. For these 26 primary studies, some variables were coded into Microsoft Excel. The coding scheme for the coded variables is given in Table S1 in Supplementary File 1. Primary studies coded as in the file were coded by two coders. To provide evidence of intercoder reliability, percent agreement was calculated according to Miles and Huberman's (1994) formula. Since all codings of the two coders were the same, the agreement percentage was found to be 100%. The descriptive characteristics of the primary studies included in the meta-analysis after coding are also presented in Table S2 in Supplementary File 1. Additionally, reference information for all included primary studies is available at Supplementary File 2.

Data analysis

In the current study, analyzes were conducted for both the 21-item (the original AIAS by Wang and Wang, 2022) and 16-item (the adapted AIAS version by Akkaya et al., 2021) versions of the scale. Meta-analytic reliability generalization was made for the overall scale and subscales in the original AIAS by Wang and Wang, and for the overall scale and Learning and Job Replacement subscales in the adapted AIAS version by Akkaya et al.. Since the number of items in the other two subscales (AI Configuration and Sociotechnical Blindness) in the adapted AIAS version by Akkaya et al. are same with the original AIAS by Wang and Wang, they were included in the analysis with the original version.

In this study, Cronbach's alpha coefficients in the primary studies were combined according to Vacha-Haase's (1998) meta-analytic reliability generalization. Bonett's transformation method (Bonett, 2002), which is considered to be theoretically more suitable for Cronbach's α coefficient (López-Ibáñez et al., 2024), was used as the transformation method. While the random effect model was used to obtain pooled Cronbach's α for the overall scales and subscales because heterogeneity statistics values and populations are different, the mixed effect model was used in moderator analyses. The inverse variance method was used as the weighting method. Also, the Restricted Maximum Likelihood method was used to estimate the variance between studies.

In the current research, firstly, publication bias was examined. For this purpose, Rosenthal's fail- safe N, Begg and Mazumdar's rank correlations, Egger regression intercept and Duval-Tweedie's trim and fill method based on funnel plot values were examined. In order to evaluate heterogeneity, Q, I^2 and τ^2 values were examined. In line with the heterogeneity results, moderator analysis was carried out on how the reliability values of the overall scale and subscales changed according to the determined categorical variables. Likert type, scale language, scale version, country, study

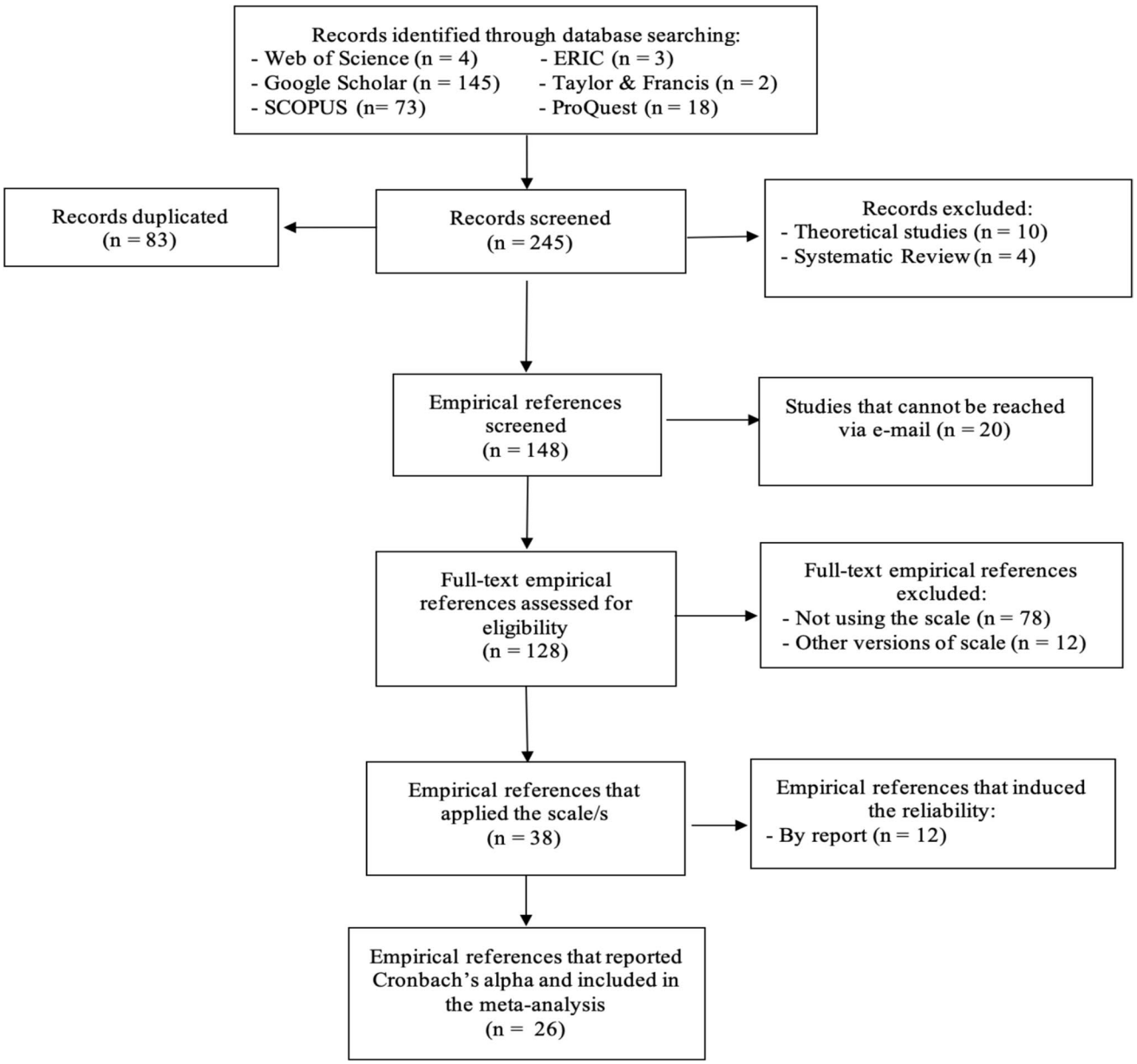


Fig. 1 REGEMA flowchart

group, research type and study field were treated as categorical variables. Analyzes were carried out by determining the variables in the overall scale and subscales in order to ensure that there were a sufficient number of studies in the subgroups created according to the independent variables considered. Also, meta-regression analysis was conducted to determine how the reliability value would change according to continuous variables in the study. In general, meta-regression is not recommended if there are fewer than ten primary studies (Higgins & Green, 2011). For this reason, only ratio of women, one of the continuous variables, was examined in this study. All analyses were conducted in the R metafor package (Viechtbauer, 2010), and the R ggplot2 package (Wickham, 2016) was used for forest plots. R scripts and data as a file was presented at the link: https://osf.io/dc6zf/?view_only=f1a317c0257a41f18eaa8c39bb2896b7

Results

In the study, publication bias, induction rate, pooled Cronbach's alpha results, heterogeneity, moderator analysis and meta regression results were examined within the meta-analytic reliability generalization analysis of reliability coefficients for AIAS.

Publication bias

Publication bias results were analyzed and are shown in Table 1 for reliability generalizations.

When Table 1 is examined, Rosenthal's fail-safe N values (Rosenthal, 1979) being higher than 5 k + 10 (k = number of studies) indicate that there is no publication bias. Kendall's τ and Egger's intercept values are not significant. In addition, only in the Learning and AI Configuration subscales of the original AIAS by Wang and Wang, it was determined that 2 and 3 studies should be added for the funnel plot to be symmetrical, respectively. However, when all values are interpreted together, it can be stated that there is no publication bias for both versions of the scales and the subscales considered.

Induction rate

According to Vacha-Haase et al. (2000), the induction rate was calculated in this study. As seen in Fig. 1, there are 40 studies using one of the scale versions or its subscales. It was observed that 12 of these 40 studies did not report reliability coefficients. Based on this, the induction rate was found to be 30%.

Pooled Cronbach's alpha results

Table 2 shows the pooled Cronbach's alpha values for the overall scale and subscales.

As shown in Table 2, all reliability values obtained are significant. In the original AIAS by Wang and Wang, the highest reliability value was obtained for the overall scale.

Table 1 Publication bias for overall scale and subscales

	Fail-safe *N*	Rank Correlation Test		Egger Regression Test		Adjusted studies
		Kendall τ	*p*	*t*	*p*	
The original AIAS by Wang and Wang						
Overall Scale	118,250	0.287	0.137	1.418	0.180	0
Learning	36,970	−0.067	0.862	0.088	0.932	2 (R)
Job Replacement	23,521	0.056	0.920	−0.268	0.796	0
Sociotechnical Blindness	28,882	−0.154	0.510	−0.450	0.662	0
AI Configuration	41,433	0.231	0.306	1.046	0.318	3 (L)
The adapted AIAS version by Akkaya et al.						
Overall Scale	27,768	−0.244	0.381	−0.728	0.487	0
Learning	6898	−0.600	0.233	−1.948	0.147	0
Job Replacement	3060	0.000	1.000	0.975	0.402	0

The original AIAS by Wang and Wang: 21 items and four subscale version

The adapted AIAS version by Akkaya et al.: 16 items and four subscale version

Fail-safe N: Number of missing studies to be added to bring the p-value statistically insignificant according to Rosenthal Failsafe N method

Adjusted Studies: Estimated number of missing studies on the right (R) or left (L) side according to the Duval-Tweedie trim and fill method

Table 2 Pooled Cronbach's Alpha of Overall scale and subscales for each version and, results of heterogeneity

	k	α	95% CI [LL_α-UL_α]	Z	*Q*	I^2 (%)	τ^2	95% CI [$LL\tau^2$ -$UL\tau^2$]
The original AIAS by Wang and Wang								
Overall Scale	15	0.949***	[0.944–0.963]	29.376	332.210***	96.10	0.158	[0.036–0.280]
Learning	10	0.930***	[0.911–0.951]	17.515	217.820***	96.99	0.232	[0.009–0.455]
Job Replacement	9	0.911***	[0.883–0.938]	15.334	147.985***	96.69	0.223	[−0.004–0.450]
Sociotechnical Blindness	13	0.877***	[0.854–0.903]	20.535	137.508***	93.07	0.128	[0.016–0.240]
AI Configuration	13	0.921***	[0.909–0.954]	15.914	310.536***	97.20	0.372	[0.064–0.680]
The adapted AIAS version by Akkaya et al.								
Overall Scale	10	0.878***	[0.844–0.935]	10.251	741.835***	98.66	0.489	[0.024–0.953]
Learning	5	0.932***	[0.896–0.949]	14.347	58.929***	94.14	0.155	[−0.076–0.386]
Job Replacement	5	0.820***	[0.785–0.893]	10.528	80.709***	93.51	0.148	[−0.075–0.371]

***$p < 0.0001$, *k*: Number of studies, α: Pooled Cronbach's Alpha Coefficient, LL_α: Lower Limit of Pooled Cronbach's Alpha Coefficient for 95% Confidence Interval, UL_α: Upper Limit of Pooled Cronbach's Alpha Coefficient for 95% Confidence Interval, *Q* and I^2: Heterogeneity Statistics and τ^2: Between-Study Variance, [$LL\tau^2$]: Lower Limit of τ^2 for 95% Confidence Interval and, [$UL\tau^2$]: Upper Limit of τ^2 for 95% Confidence Interval

The lowest reliability value was obtained for the Sociotechnical Blindness subscale. For the adapted AIAS version by Akkaya et al., the highest reliability value was obtained from Learning subscale and the lowest reliability value was estimated from Job Replacement subscale. The forest plot for the overall scale of the original AIAS by Wang and Wang andthe adapted AIAS version by Akkaya et al. is given in the Fig. 2 . When both versions of the scales were compared, higher reliability values were obtained in the original AIAS by Wang and Wang for the overall scale and subscales.

Assessing heterogeneity

To assess heterogeneity, Q, I^2 and τ^2 values are presented in Table 2. The Q values (Cochran, 1954) are significant for both versions and subscales of the AIAS. When I^2 values are analyzed, it can be stated that there is high heterogeneity (Higgins et al., 2003) in the overall scale versions and their subscales. As another criterion for examining heterogeneity, τ^2 values were examined and it was determined that all values were different from 0. In other words, it can be stated that there is variance between studies. When all the values obtained regarding heterogeneity are analyzed, it can be stated that there is heterogeneity, that is, the variance between the studies considered within the scope of the research. For this reason, analysis results based on the random effects model are reported for the overall scale and subscales in both versions of the scales. In addition to the statistical evidence of heterogeneity, variables such as having different versions of the scale and being applied in different populations also evidenced heterogeneity in the theoretical context and this led to the use of a random effect model.

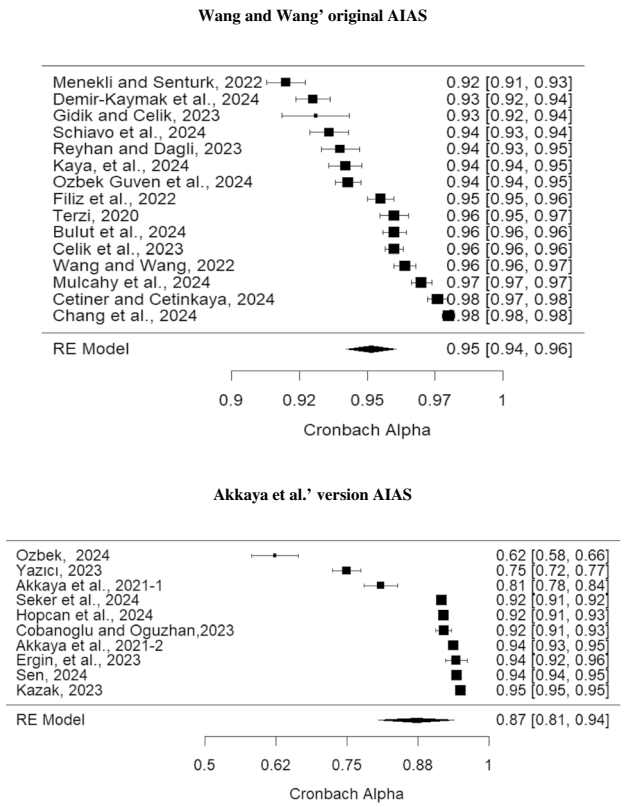


Fig. 2 Forest plots for both AIAS versions

Moderator analysis results

Table 3 presents the results of analog to ANOVA for the overall and subscales of the original AIAS by Wang and Wang and the adapted AIAS version by Akkaya et al. according to the categories determined within the scope of the research.

When Table 3 is examined, it is seen that all reliability values are significant in the original AIAS by Wang and Wang. According to the ANOVA results, it was determined that the reliability value of the overall scale showed a significant difference in the countries and study fields. In the country category, the reliability value obtained in the China-Taiwan subgroup was higher than the Türkiye subgroup. In the study field category, the highest Cronbach's alpha was obtained in the management subgroup, while the lowest Cronbach's alpha was obtained in the medicine subgroup. It was determined that the Cronbach's alpha of the overall scale did not show a significant difference in Likert type, scale language, study group and research type categories. In the subscales, no significant difference was observed in the categories considered. When all reliability values obtained in the AIAS version adapted by Akkaya et al. are examined in Table 3, they are significant for the overall scale and the subscales. According to the ANOVA results, the reliability value of the overall scale does not show a significant difference in the study group, research type and study field categories. In the subscale of Job Replacement, while the difference in the reliability value was statistically significant in the study group, it was determined that the difference was not statistically significant according to the research type. In the subscale of Learning, it was determined that the difference was not statistically significant according to the research type and study group.

Table 4 shows the Meta regression findings obtained according to the ratio of women variable.

When Table 4 is examined, only the α coefficient of the Job Replacement subscale of the original AIAS by Wang and Wang shows a significant difference regarding the continuous variable of women ratio. In the Job Replacement subscale, the α coefficient of the women ratio variable explains 3.37% of the variance ($R^2 = 0.03$). In addition, the fact that Q_E is significant in all overall scales and subscales

Table 3 Moderator analysis results for the original AIAS by Wang and Wang , the adapted AIAS version by Akkaya et al. and theirs subscales

	k	α [LLα-ULα]	*Q*	*df*	*p*
The original AIAS by Wang and Wang					
Overall Scale					
Likert Type					
5-point	5	0.965*** [0.951–0.975]	3.253	1	0.071
7-point	10	0.948***[0.935–0.959]			
Scale Language					
European Languages	3	0.959*** [0.939–0.972]	0.792	1	0.373
Turkish	11	0.950*** [0.938–0.959]			
Country					
China-Taiwan	2	0.973*** [0.956–0.984]	5.446	1	0.019***
Türkiye	11	0.950*** [0.938–0.959]			
Study Group					
Student	3	0.938*** [0.904–0.960]	2.422	1	0.120
Adult	12	0.958*** [0.948–0.966]			
Research Type					
Scale Development/Adaptation	3	0.965*** [0.945–0.978]	1.576	1	0.209
Survey Model	12	0.951*** [0.939–0.961]			
Study Field					
Education & Psychology	5	0.950*** [0.937–0.961]	19.854	3	0.000***
Management	3	0.976*** [0.967–0.982]			
Medicine	5	0.943*** [0.927–0.955]			
Sport Sciences	2	0.948*** [0.924–0.965]			
Learning					
Likert Type					
5-point	2	0.949*** [0.892–0.976]	0.588	1	0.443
7-point	7	0.929*** [0.895–0.953]			
Scale Language					
European Languages	2	0.954*** [0.907–0.978]	1.196	1	0.274
Turkish	7	0.928*** [0.895–0.951]			
Country					
China-Taiwan	2	0.955*** [0.909–0.978]	1.283	1	0.257
Türkiye	7	0.928*** [0.895–0.951]			
Research Type					
Scale Development/Adaptation	2	0.947*** [0.892–0.973]	0.429	1	0.512
Survey Model	8	0.930*** [0.901–0.951]			
Study Field					
Education & Psychology	5	0.934*** [0.897–0.958]	2.129	2	0.345
Management	2	0.953*** [0.903–0.977]			
Sport Sciences	2	0.900*** [0.793–0.951]			
Job Replacement					
Likert Type					
5-point	2	0.910*** [0.803–0.959]	0.032	1	0.857
7-point	6	0.917*** [0.870–0.947]			
Scale Language					
European Languages	2	0.891*** [0.782–0.946]	0.578	1	0.447
Turkish	7	0.920*** [0.884–0.945]			
Research Type					
Scale Development/Adaptation	2	0.935*** [0.872–0.967]	0.845	1	0.358
Survey Model	7	0.907*** [0.866–0.936]			
Study Field					
Education & Psychology	4	0.920*** [0.869–0.951]	1.997	2	0.368
Management	2	0.927*** [0.853–0.964]			
Sport Sciences	2	0.862*** [0.722–0.931]			
Sociotechnical Blindness					
Likert Type					

Table 3 (continued)

	k	α [LLα-ULα]	*Q*	*df*	*p*
5-point	6	0.880*** [0.835–0.913]	0.003	1	0.958
7-point	7	0.881*** [0.841–0.911]			
Scale Version					
21 items	8	0.888*** [0.853–0.914]	0.522	1	0.470
16 items	5	0.868*** [0.815–0.906]			
Scale Language					
European Languages	2	0.872*** [0.778–0.926]	0.076	1	0.783
Turkish	11	0.882*** [0.851–0.907]			
Study Group					
Student	3	0.873*** [0.802–0.919]	0.085	1	0.770
Adult	10	0.883*** [0.850–0.908]			
Research Type					
Scale Development/Adaptation	4	0.895*** [0.846–0.928]	0.613	1	0.434
Survey Model	9	0.874*** [0.837–0.902]			
Study Field					
Education & Psychology	5	0.875*** [0.822–0.913]	1.356	3	0.716
Management	4	0.900*** [0.850–0.933]			
Medicine	2	0.877*** [0.783–0.931]			
Sport Sciences	2	0.851*** [0.735–0.916]			
AI Configuration					
Likert Type					
5-point	6	0.946*** [0.910–0.967]	0.885	1	0.347
7-point	7	0.925*** [0.880–0.953]			
Scale Version					
21 items	8	0.937*** [0.901–0.960]	0.048	1	0.826
16 items	5	0.932*** [0.879–0.962]			
Scale Language					
European Languages	2	0.947*** [0.870–0.978]	0.225	1	0.635
Turkish	11	0.933*** [0.902–0.954]			
Study Group					
Student	3	0.942*** [0.880–0.972]	0.126	1	0.723
Adult	10	0.933*** [0.900–0.955]			
Research Type					
Scale Development/Adaptation	4	0.956*** [0.922–0.975]	2.497	1	0.114
Survey Model	9	0.923*** [0.887–0.948]			
Study Field					
Education & Psychology	5	0.939*** [0.901–0.963]	6.189	3	0.103
Management	4	0.951*** [0.915–0.972]			
Medicine	2	0.945*** [0.881–0.975]			
Sport Sciences	2	0.844*** [0.661–0.928]			
The adapted AIAS version by Akkaya et al.					
Overall Scale					
Study Group					
Student	4	0.875*** [0.747–0.938]	0.606	1	0.436
Adult	6	0.913*** [0.845–0.951]			
Research Type					
Scale Development/Adaptation	2	0.890*** [0.681–0.962]	0.004	1	0.950
Survey Model	7	0.895*** [0.814–0.940]			
Study Field					
Education & Psychology	2	0.856*** [0.602–0.950]	1.054	2	0.590
Management	5	0.883*** [0.776–0.939]			
Medicine	2	0.932*** [0.805–0.976]			
Learning					
Study Group					

Table 3 (continued)

	k	α [LLα-ULα]	*Q*	*df*	*p*
Student	3	0.925*** [0.873–0.956]	0.009	1	0.924
Adult	2	0.929*** [0.862–0.963]			
Research Type					
Scale Development/Adaptation	2	0.923*** [0.852–0.960]	0.035	1	0.852
Survey Model	3	0.929*** [0.879–0.958]			
Job Replacement					
Study Group					
Student	3	0.883*** [0.855–0.906]	14.118	1	0.000***
Adult	2	0.774*** [0.704–0.827]			
Research Type					
Scale Development/Adaptation	2	0.890*** [0.829–0.929]	3.269	1	0.071
Survey Model	3	0.814*** [0.734–0.870]			

***$p<0.0001$, *k*: Number of studies, α: Pooled Cronbach's Alpha Coefficient, LL_{α}: Lower Limit of Pooled Cronbach's Alpha Coefficient for 95% Confidence Interval, UL_{α}: Upper Limit of Pooled Cronbach's Alpha Coefficient for 95% Confidence Interval, *Q*: The statistic for test of moderators, *df*: degree of freedom (number of categories—1)

Table 4 Meta regression results for ratio of women

	k	β [LL_{β}—UL_{β}]	z	p	Q_E	R^2
The original AIAS by Wang and Wang						
Overall Scale	15	−0.503 [−1.363–0.357]	−1.146	0.252	254.717***	0.01
Learning	10	0.482 [−1.081–2.045]	0.604	0.546	217.760***	0.00
Job Replacement	9	1.339* [0.090–2.588]	2.101	0.036	126.758	0.03
Sociotechnical Blindness	13	0.276 [−0.735–1.286]	0.535	0.593	136.336***	0.00
AI Configuration	13	1.273 [−0.231–2.777]	1.658	0.097	291.552***	0.01
The adapted AIAS version by Akkaya et al.						
Overall Scale	10	0.865 [−1.825–3.555]	0.630	0.529	741.804***	0.00

*$p<0.05$, ***$p<0.0001$, β, Slope; Q_E, Weighted error sum of squares and R^2, Proportion of variance explained

indicates that variables other than the proportion of women can explain the variability in α coefficient.

Discussion

As a revolutionary technology, AI technology is developing rapidly. In addition to its benefits, AI has been known to cause AI anxiety for individuals (Kaya et al., 2024). Considering AI anxiety, provide more accurate use of AI technology. So, determining the level of AI anxiety is crucial. Therefore, it is important to use a tool with strong psychometric properties that can measure AI anxiety. As AI anxiety is a global phenomenon, it is important to determine their psychometric properties by combining measurements obtained from samples from different populations. AI anxiety can be measured with Artificial Intelligence Anxiety Scale (AIAS) and it was considered two versions of the scale in this study. The first version of the scale is Wang and Wang's 21-item version and another version is Akkaya et al.'s 16-item version. Both scale versions consist of subscales with the identical name. However, the number of items in the Job Replacement and Learning subscales differs in the two versions.

The first goal of this research was to conduct an RG analysis to synthesize the score reliability estimates for the original AIAS by Wang and Wang and the adapted AIAS version by Akkaya et al.. When the calculated overall internal consistency score reliability estimates were examined, it is higher for the the original AIAS by Wang and Wang compared tothe adapted AIAS version by Akkaya et al.. The lower reliability estimate of the original AIAS by Wang and Wang is higher than the upper reliability estimate of the adapted AIAS version by Akkaya et al.. In other words, the pooled reliability estimates of the two scale versions are different from each other in terms of confidence intervals. The differences between the pooled reliability estimates is 0.07 between the original AIAS by Wang and Wang and the adapted AIAS version by Akkaya et al.. This difference is in favor of the original AIAS by Wang and Wang. This may be due to the fact that internal consistency estimates are affected by the test length (Cortina, 1993), and it may also be attributed to more measurement error to the adapted AIAS version by Akkaya et al.. As Şen's (2021) reliability generalization study, the long version of the test tended to have a higher Cronbach's alpha value than the short version. On the other hand, both version of scales and subscales considered

within the scope of the research, produced scores that were adequately reliable based on the pooled Cronbach's α. When the subscales that differ from each other in the two scale versions are examined, the pooled Cronbach's alpha values for the Learning subscale are quite similar. However, the pooled Cronbach's alpha for Job Replacement subscale is higher in favor of the original AIAS by Wang and Wang.

Secondly, the goal of this study was to investigate moderating variables that could explain the variations in score reliability estimates for the the original AIAS by Wang and Wang and the adapted AIAS version by Akkaya et al.. For the original AIAS by Wang and Wang, six categorical (likert type, scale version, scale language, study group, research type, study field); for the adapted AIAS version by Akkaya et al. three categorical (study group, research type and study field) and for both of them, one continuous variable (ratio of women) was tested as moderator variables. In meta-analyses studies, moderator variables are used to examine the heterogeneity, that is, the difference in the Cronbach's alpha in the studies included in the analysis. All reliability values obtained in the current study were above the acceptable value. However, revealing the reason for the differences between the studies is important in terms of the measures that can be taken regarding the accuracy of the results obtained with the scale. The lack of differences in the moderator analyses may be attributed to the fact that some studies did not present data on the variables considered. Or the variables determined within the scope of the research may be insufficient to explain the heterogeneity between the studies. While no moderating effects were found for the overall scale of the adapted AIAS version by Akkaya et al., country and study field are the significant moderator variable for the overall scale of the original AIAS by Wang and Wang. For the country variable, the reliability estimates were higher in China-Taiwan than in Türkiye. This may be due to the fact that the scale is clearer in the culture in which it was developed than in other cultures. As in reliability generalization studies that include scale adaptation studies, it has been observed that the country variable tends to be higher in the culture in which the scale was developed (Tümtürk, 2024; Yiğiter et al., 2024). For the study field variable, the reliability estimates were from highest to smallest is as follows: management, medicine, education and psychology and sports sciences. The Cronbach's α reliability value obtained from the studies conducted in the field of education and psychology is higher than in the field of medicine. This may be due to the fact that the number of studies conducted in the field of education and psychology is higher than in the field of medicine. The increase in the heterogeneity of the sample may also be due to the increase in the number of studies. In addition, the fact that the groups working in the field of medicine are more homogeneous may have increased the heterogeneity of the sample. While no moderating effects were found for the subscale of the original AIAS by Wang and Wang, the study group is the significant moderator variable for Job Replacement subscale of the adapted AIAS version by Akkaya et al.. Some RG studies have shown that the study group is a minor source of variability in score reliability estimates (Davis et al., 2024; Şahin & Yıldırım, 2024), in this study, on the contrary, students have more reliable Job Replacement scores than adults. It should be noted, however, that the reliability for the Job Replacement subscale of the adapted AIAS version by Akkaya et al. for both groups is quite high. When the results of the continious moderator analysis are examined, it is seen that ratio of women has no significant effect except for the Job Replacement subscale in the original AIAS by Wang and Wang. In other words, it can be said that an increase in the ratio of women increases the reliability estimation value for this subscale. Although the women's scores tend to show higher internal consistency, this continuous variable explained 3.37% of the variance in the Job Replacement subscale of the original AIAS by Wang and Wang. This result suggests that gender distribution, as operationalized in this study, has a limited effect on the reliability for this subscale. Similar to the results of this study, there are studies in the literature that found that an increase in the proportion of women increases the reliability estimation value (Beretvas et al., 2002; Lenz et al., 2024).

In conclusion, the results of this research support the internal consistency of the both version of scales and their use within differing sample characteristics'. Especially for the AIAS, which Wang and Wang (2020) demonstrated to provide satisfactory content validity, discriminant validity, convergent validity, and nomological validity, in current study also revealed that it provides satisfactory reliability. So based on the findings of current study, AIAS will be useful for researchers who want to conduct research on anxiety towards AI, practitioners who want to make accurate decisions implementing AI technologies, and educators who want to employ the in learning environments with AI. Finally, the results of this research support the internal consistency of the both version of scales and their use within differing sample characteristics'.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s12144-025-08736-5.

Author contributions All authors contributed to the study conception and design equally. Material preparation and data collection were conducted by Tuba Gündüz and Yıldız Yıldırım. Formal analysis was performed Yıldız Yıldırım. The first draft of the manuscript was written by Yıldız Yıldırım, Tuba Gündüz and Melek Gülşah Şahin. The manuscript was edited in the line with reviewer opinion by Yıldız Yıldırım, Tuba Gündüz and Melek Gülşah Şahin. All authors read and approved the final manuscript.

Funding Open access funding provided by the Scientific and Technological Research Council of Türkiye (TÜBİTAK). No specific grant from a public, private, or nonprofit organization was obtained for this study.

Data availability We have shared the link to the data and R scripts in the Manuscript.

Declarations

Ethics approval Not applicable.

Informed consent This meta-analytic reliability generalization study did not involve human participants or data directly. All data come from previous primary studies. Therefore, informed consent is not required.

Conflict of interest The authors state that there is no conflict of interest.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

References

Akkaya, B., Özkan, A., & Özkan, H. (2021). Yapay zekâ kaygı (YZK) ölçeği: Türkçeye uyarlama, geçerlik ve güvenirlik çalışması. *Alanya Akademik Bakış, 5*(2), 1125–1146. https://doi.org/10.29023/alanyaakademik.833668

Batı, K., & Irmak, Ş. (2024). A reliability generalization study of the STEM-CIS scale: Exploring moderator effects. *Current Psychology, 43*(46), 35453–35470. https://doi.org/10.1007/s12144-024-07036-8

Beretvas, S. N., Meyers, J. L., & Leite, W. L. (2002). A reliability generalization study of the Marlowe-Crowne Social Desirability Scale. *Educational and Psychological Measurement, 62*(4), 570–589. https://doi.org/10.1177/0013164402062004003

Blázquez-Rincón, D., Durán, J. I., & Botella, J. (2022). The fear of COVID-19 scale: A reliability generalization meta-analysis. *Assessment, 29*(5), 940–948. https://doi.org/10.1177/1073191121994164

Bonett, D. G. (2002). Sample size requirements for testing and estimating coefficient alpha. *Journal of Educational and Behavioral Statistics, 27*(4), 335–340. https://doi.org/10.3102/10769986027004335

Chan, K. S., & Zary, N. (2019). Applications and challenges of implementing artificial intelligence in medical education: Integrative review. *JMIR Medical Education, 5*(1), Article e13930. https://doi.org/10.2196/13930

Colosimo, B. M., del Castillo, E., Jones-Farmer, L. A., & Paynabar, K. (2021). Artificial intelligence and statistics for quality technology: An introduction to the special issue. *Journal of Quality Technology, 53*(5), 443–453. https://doi.org/10.1080/00224065.2021.1987806

Comer, J. (2023). The psychological fears associated with AI. *Psychology Today.* https://www.psychologytoday.com/us/blog/beyond-stress-and-burnout/202307/the-psychological-fears-associated-with-ai

Cortina, J. M. (1993). What is coefficient alpha? An examination of theory and applications. *Journal of Applied Psychology, 78*, 98–104. https://doi.org/10.1037/0021-9010.78.1.98

Dai, Y., Chai, C.-S., Lin, P.-Y., Jong, M.S.-Y., Guo, Y., & Qin, J. (2020). Promoting students' well-being by developing their readiness for the artificial intelligence age. *Sustainability, 12*(16), Article 6597. https://doi.org/10.3390/su12166597

Eser, M. T., & Doğan, N. (2023). Life satisfaction scale: A meta-analytic reliability generalization study in Turkey sample. *Türk Psikolojik Danışma ve Rehberlik Dergisi, 13*(69), 224–239. https://doi.org/10.17066/tpdrd.1223320mn

Gull, A., Dilawar, S., & Sher, F. (2023). Data-driven artificial intelligence at the crossroads: Investigating the role of affective job insecurity in the relationship between artificial intelligence identity threat and employee well-being. *The Asian Bulletin of Big Data Management, 3*(1), 18–34. https://doi.org/10.62019/abbdm.v3i1.35

Higgins, J. P., & Green, S. (2011). *Cochrane handbook for systematic reviews of interventions (Version 5.1.0) [Handbook, updated March 2011]*. The Cochrane Collaboration.

Johnson, D. G., & Verdicchio, M. (2017). AI anxiety. *Journal of the Association for Information Science and Technology, 68*(9), 2267–2270. https://doi.org/10.1002/asi.23867

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

Kim, J. J. H., Kadkol, S., Solomon, I., Yeh, H., Soh, J. Y., Nguyen, T. M., Choi, J. Y., Lee, S., Srivatsa, A. V., Nahass, G. R., & Ajilore, O. A. (2023). AI anxiety: A comprehensive analysis of psychological factors and interventions [Preprint]. *SSRN*. https://doi.org/10.2139/ssrn.4573394

Kwak, Y., Ahn, J. W., & Seo, Y. H. (2022). Influence of AI ethics awareness, attitude, anxiety, and self-efficacy on nursing students' behavioral intentions. *BMC Nursing, 21*(1), Article 267. https://doi.org/10.1186/s12912-022-01048-0

Lenz, A. S., Smith, C., & Meegan, A. (2024). A reliability generalization meta-analysis of scores on the Professional Quality of Life (ProQOL) Scale across sample characteristics. *Measurement and Evaluation in Counseling and Development, 57*(4), 403–417. https://doi.org/10.1080/07481756.2024.2320154

Li, J., & Huang, J. S. (2020). Dimensions of artificial intelligence anxiety based on the integrated fear acquisition theory. *Technology in Society, 63*, Article 101410. https://doi.org/10.1016/j.techsoc.2020.101410

López-Ibáñez, C., López-Nicolás, R., Blázquez-Rincón, D. M., & Sánchez-Meca, J. (2024). Reliability generalization meta-analysis: Comparing different statistical methods. *Current Psychology, 43*, 18275–18293. https://doi.org/10.1007/s12144-023-05604-y

López-Pina, J. A., Sánchez-Meca, J., López-López, J. A., Marín-Martínez, F., Núñez-Núñez, R. M., Rosa-Alcázar, A. I., Gómez-Conesa, A., & Ferrer-Requena, J. (2015). The Yale-Brown obsessive compulsive scale: A reliability generalization meta-analysis. *Assessment, 22*(5), 619–628. https://doi.org/10.1177/1073191114551954

McCarthy, J. (2007). What is artificial intelligence? *The Journal for Quality and Participation, 29*(4), 4–7.

Miles, M. B., & Huberman, A. M. (1994). *Qualitative data analysis: An expanded sourcebook.* Sage.

Neudert, L. M., Knuutila, A., & Howard, P. N. (2020). *Global attitudes towards AI, machine learning & automated decision making* (Working paper 2020.10). Oxford Commission on AI & Good Governance. https://oxcaigg.oii.ox.ac.uk/

Orhan, A. (2022). California critical thinking disposition inventory: Reliability generalization meta-analysis. *Journal of Psychoeducational Assessment, 40*(2), 202–220. https://doi.org/10.1177/07342829211048962

Özdemir, V., Yıldırım, Y., & Tan, Ş. (2020). A meta-analytic reliability generalization study of the Oxford Happiness Scale in Turkish sample. *Journal of Measurement and Evaluation in Education and Psychology, 11*(4), 374–404. https://doi.org/10.21031/epod.766266

Pektaş, T. (2024). The impact of artificial intelligence anxiety on employees: A comprehensive review of psychological and organizational dynamics. *International Journal of Education Technology and Scientific Researches, 9*(26), 194–206. https://doi.org/10.35826/ijetsar.733

Rayhan, S., & Rayhan, A. (2023). The psychological impact of AI: Adapting to a world of smart machines. ResearchGate. https://doi.org/10.13140/RG.2.2.29902.64329/1

Rubio-Aparicio, M., Badenes-Ribera, L., Sánchez-Meca, J., Fabris, M. A., & Longobardi, C. (2020). A reliability generalization meta-analysis of self-report measures of muscle dysmorphia. *Clinical Psychology: Science and Practice, 27*(1), Article e12303.

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

Sánchez-Meca, J., Marín-Martínez, F., López-López, J. A., Núñez-Núñez, R. M., Rubio-Aparicio, M., López-García, J. J., López-Pina, J. A., Blázquez-Rincón, D., López-Ibáñez, C., & López-Nicolás, R. (2021). Improving the reporting quality of reliability generalization meta-analyses: The REGEMA checklist. *Research Synthesis Methods, 12*(4), 516–536. https://doi.org/10.1002/jrsm.1487

Şen, S. (2021). A reliability generalization meta-analysis of Runco Ideational Behavior Scale. *Creativity Research Journal, 34*(2), 178–194. https://doi.org/10.1080/10400419.2021.1960719

Terzi, R. (2020). An adaptation of Artificial Intelligence Anxiety Scale into Turkish: Reliability and validity study. *International Online Journal of Education and Teaching, 7*(4), 1501–1515.

Tugay, B., & Tugay, R. (2019). Uluslararası sistemin geleceğini yapay zekâ üzerinden analiz etmek. *Journal of Academic Value Studies, 5*(3), 376–384.

Tümtürk, E. (2024). *Çocukların Sınav Kaygısı Ölçeğinin güvenirlik genelleme meta-analiz çalışması.* [Master Thesis]. Harran Üniversitesi, Sosyal Bilimler Enstitüsü.

Vacha-Haase, T. (1998). Reliability generalization: Exploring variance in measurement error affecting score reliability across studies. *Educational and Psychological Measurement, 58*(1), 6–20. https://doi.org/10.1177/0013164498058001002

Vacha-Haase, T., Kogan, L. R., & Thompson, B. (2000). Sample compositions and variabilities in published studies versus those in test manuals: Validity of score reliability inductions. *Educational and Psychological Measurement, 60*(4), 509–522. https://doi.org/10.1177/00131640021970682

Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor package. *Journal of Statistical Software, 36*(3), 1–48. https://doi.org/10.18637/jss.v036.i03

Wang, K., Ruan, Q., Zhang, X., Fu, C., & Duan, B. (2024). Pre-service teachers' GenAI anxiety, technology self-efficacy, and TPACK: Their structural relations with behavioral intention to design GenAI-assisted teaching. *Behavioral Science, 14*(5), Article 373. https://doi.org/10.3390/bs14050373

Wang, Y. Y., & Wang, Y. S. (2022). Development and validation of an artificial intelligence anxiety scale: An initial application in predicting motivated learning behavior. *Interactive Learning Environments, 30*(4), 619–634. https://doi.org/10.1080/10494820.2019.1674887

Wang, Y. M., Wei, C. L., Lin, H. H., Wang, S. C., & Wang, Y. S. (2022). What drives students' AI learning behavior: A perspective of AI anxiety. *Interactive Learning Environments, 32*(6), 2584–2600. https://doi.org/10.1080/10494820.2022.2153147

Wickham, H. (2016). *ggplot2: Elegant graphics for data analysis*. Springer.

Yiğiter, M. S., Eser, M. T., & Doğan, N. (2024). Reliability generalization of social appearance anxiety scale: A meta analysis study. *Hacettepe University Journal of Education, 39*(1), 97–112. https://doi.org/10.16986/HUJE.2024.512

Yıldırım, Y. (2023). Meta-analizde gri literatür [Gray literature in meta-analysis]. In E. Bay & O. Zahal (Eds.), *Eğitim bilimleri alanında akademik çalışmalar [Academic studies in educational science]* (pp. 329–344). Gece Kitaplığı.

Yörük, S., & Şen, S. (2023). A reliability generalization meta-analysis of the creative achievement questionnaire. *Creativity Research Journal, 35*(4), 714–729.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.