**International Journal of Assessment Tools in Education**

2024, Vol. 11, No. 2, 303–319

https://doi.org/10.21449/ijate.1369023

journal homepage: https://dergipark.org.tr/en/pub/ijate *Research Article*

# The general attitudes towards artificial intelligence (GAAIS): A meta-analytic reliability generalization study

**Melek Gülşah Şahin** [1,*], **Yıldız Yıldırım** [2]

[1]Gazi University, Gazi Faculty of Education, Department of Educational Sciences, Ankara, Türkiye

[2]Aydın Adnan Menderes University, Faculty of Education, Department of Educational Sciences, Aydın, Türkiye

**ARTICLE HISTORY**

*Received: Sep. 30, 2023*

*Accepted: Mar. 22, 2024*

**Keywords:**

General attitudes towards artificial intelligence scale,
Reliability generalization,
Cronbach's α,
Mc Donald's ω,
Moderator analysis.

**Abstract:** This study aims to generalize the reliability of the GAAIS, which is known to perform valid and reliable measurements, is frequently used in the literature, aims to measure one of today's popular topics, and is one of the first examples developed in the field. Within the meta-analytic reliability generalization study, moderator analyses were also conducted on some categorical and continuous variables. Cronbach's α values for the overall scale and the positive and negative subscales, and McDonald's ω coefficients for positive and negative subscales were generalized. Google Scholar, WOS, Taylor & Francis, Science Direct, and EBSCO databases were searched to obtain primary studies. As a result of the screening, 132 studies were found, and these studies were reviewed according to the inclusion criteria. Reliability coefficients obtained from 19 studies that met the criteria were included in the meta-analysis. While meta-analytic reliability generalization was performed according to the random effects model, moderator analyses were performed according to the mixed effect model based on both categorical variables and continuous variables. As a result of the research pooled, Cronbach's α was 0.881, 0.828, and 0.863 for total, the negative, and positive subscales respectively. Also, McDonald's ω was 0.873 and 0.923 for negative and positive subscales respectively. It was found that there were no significant differences between the reliability coefficients for all categorical variables. On the other hand, all continuous moderator variables (mean age, standard deviation age, and rate of female) had a significant effect.

## 1. INTRODUCTION

In everyday life, applications related to artificial intelligence are encountered or used almost daily. Reasons such as the fact that computers play an essential role in our lives and that their use increases due to the convenience they bring and the different experiences they offer every day and that they eliminate the problem of space and time in accessing information, provide an understanding of the popularity of artificial intelligence. There are many studies related to artificial intelligence in many fields. When the keyword "artificial intelligence" is searched in Google Scholar, 3.490.000 research studies are found. Especially ChatGPT, which is one of the most important AI applications recently, maintains its popularity in all fields.

*CONTACT: Melek Gülşah Şahin ✉ mgulsahsahin@gazi.edu.tr Gazi University, Gazi Faculty of Education, Department of Educational Sciences, Ankara, Türkiye.



**e-ISSN: 2148-7456**

Artificial intelligence is the combination of science and engineering to create intelligent computers or programs to perform tasks related to human intelligence (McCarthy, 2004). Here, human intelligence and machines interact with each other. In other words, artificial intelligence is defined as the ability of a computer to perform features such as reasoning, problem-solving, inferring, and generalizing, as in humans (Arslan, 2020). One of the most important examples of artificial intelligence is Cog and Kismet, developed in MIT laboratories in the 1990s. While Cog is an upper-body robot with visual, emotional, and kinesthetic features, Kismet is a robot head with active vision and facial expressions (Turkle et al., 2006). In addition, the most crucial feature of Kismet and Cog is that they are robots with humanoid behaviors that can communicate emotionally and socially with people. While traditional robots are equipped with applications for less communication with humans, Kismet and Cog are social robots open to sharing with humans (Breazeal, 2004). In the field of education, one of the first applications of artificial intelligence was Skinner's individualized teaching machines implemented in 1958 (Arslan, 2020). Today, artificial intelligence applications show themselves in all fields without slowing down. In addition, our accessibility to artificial intelligence is increasing day by day. SIRI, which is one of the smartphone applications, is also an AI application (Kaplan & Haenlein, 2019). In addition, many applications such as autonomous cars, virtual classrooms, face recognition, patient tracking systems, instant language translators, automation, investment tools, games, and language translations are AI applications that are constantly developing and updating themselves (Arslan, 2020; Wang et al., 2022)

Artificial intelligence applications both facilitate human life and help them gain new knowledge and experiences. Examining individuals' knowledge, experiences, attitudes, and opinions toward AI applications also contributes to the literature on the development of artificial intelligence technology. In the literature, there are studies examining individuals' attitudes toward AI applications that manifest themselves in almost all fields and that we benefit from their applications or results practically every day in our lives. For example, in the field of economics, the effect of the use of artificial intelligence in shopping on consumers' decision-making processes (Nica, 2022), the determination of attitudes toward the use of artificial intelligence in personal financial planning (Waliszewski, 2020) and in the field of health, the attitudes of dermatologists towards the use of artificial intelligence in dermatology (Polesie, 2020), the attitudes of medical students towards the use of artificial intelligence in radiology and general medicine (Pinto dos Santos et al., 2019), the attitudes and perceptions of dental students towards the use of artificial intelligence in various clinical tasks have been examined (Yuzbasioglu, 2021). It can be stated that artificial intelligence and computer technology have an important place in educational life from kindergarten to university (Kandlhofer et al., 2016). There are also studies in the field of education such as the effect of using artificial intelligence in learning environments on students' attitudes (Huang, 2018), investigating artificial intelligence anxiety and attitudes toward machine learning in pre-service teachers (Hopcan et al., 2023), and investigating university students' attitudes towards the use of SIRI in English as a foreign language (EFL) learning (Haryanto, 2019).

When the related literature is examined, it is seen that there are studies in which different measurement tools have been developed to determine attitudes toward artificial intelligence. Some of these are Attitude Towards Artificial Intelligence Scale (ATAI) (Sindermann et al., 2020), Threats of Artificial Intelligence Scale (TAI) (Kieslich et al., 2021), Negative Attitude towards Artificial Intelligence Scale (NAAIS) (Persson et al., 2021), General Attitudes Towards Artificial Intelligence Scale (GAAIS) (Schepman & Rodway, 2020; 2023), and AI Attitude Scale (AIAS-4) (Grassini, 2023). Within the scope of this research, the scales used for artificial intelligence were examined and it was determined that the most cited attitude scale was the General Attitudes Towards Artificial Intelligence Scale (GAAIS). There are 134 Google Scholar citations of the term, 113 in 2020 when GAAIS was developed, and 21 in 2023 while it has been cited a total of 53 times in the Web of Science. The other reason for selecting this scale is that validity and reliability studies have been conducted in different cultures, such as Türkiye (Kaya et al., 2022), Korea (Seo & Ahn, 2022), Finland (Bergahdl et al., 2023), and Germany (Carolus et al., 2023). The scale Schepman and Rodway (2020) developed includes 20 items and two sub-dimensions. While the positive subscale represents social and personal benefits, the negative subscale represents concerns. The developed scale was applied to 100 people (50 women and 50 men) over the age of 18 who were not students. The majority of the respondents worked in the service sector. They observed jobs from a variety of socioeconomic classes (such as cleaner, caretaker, linen assistant, sales assistant, etc.) and created 16 positive items (opportunities, benefits, and positive emotions) and 16 negative items (concerns and negative emotions) that mirrored the positive and negative themes discovered from the literature. Explanatory Factor Analysis (EFA) was conducted for the items developed. Seven items were removed from the scale based on the item correlation matrix because they showed a low correlation with other items. As a result of the EFA performed for the remaining 25 items, five items were removed because four items had factor loading values below 0.40 and 1 item had equal loading values in both dimensions, leaving 20 items. EFA was applied again to the remaining 20 items. Bartlett's Test of Sphericity $\chi^2 = 817$, $df = 190$, $p < .001$, and The Kaiser-Meyer-Olkin Measure of Sampling Adequacy (KMO MSA) was 0.86, indicating adequate sample size. In the final model, eight items were loaded on the first factor (negative views of AI), and twelve were loaded onto the second (positive attitudes towards AI). In this way, the assumptions made when the items were created about their positivity and negativity were statistically justified, providing good construct validity for the factor structure. The first and second factors accounted for 25.6% and 15.5% of the variance, respectively. The model fit measures were RMSEA = 0.0573, 90% CI [0.007, 0.068], TLI = 0.94, and the model test $\chi^2 = 182$, $df = 151$, $p = .046$, which are acceptable, namely positive attitudes towards AI ($\alpha = 0.88$) and negative attitudes towards AI ($\alpha = 0.83$).

For the validity evidence of the scale, the Technology Readiness Index scale (TRI), consisting of 18 items and four sub-dimensions, was applied to the study group and correlation and regression analyses were performed with innovation, optimism, discomfort, and insecurity sub-dimensions. The sub-dimensions of the scale were taken as independent variables, positive attitudes toward AI and negative attitudes towards AI were taken as dependent variables, and regression analysis was performed. GAAIS was reported as a valid and reliable scale when all the data were evaluated together.

Schepman and Rodway (2022), in a two-stage study (Study 1 and Study 2) that they considered the second dimension of scale development, applied the previously developed GAAIS to a sample group of 304 people to conduct CFA in Study 1 and examined its construct validity. They examined various model fit indices: $\chi^2 = 223.08$, $df = 169$, $p = 0.003$, $\chi^2/df = 1.32$, CFI = 0.987, TLI = 0.986, SRMR = 0.065, RMSEA = 0.032, 90% CI [0.019, 0.044], $p = 0.997$, suggesting an imperfect fit. In Study 1; the researchers found the standard solutions as a range of 0.310 – 0.851 for 8-item negative attitudes towards the AI subscale (GAAIS-NA); they also found that the standard solutions were in the range the range of 0.464 – 0.803 for the 12-item positive attitudes towards AI subscale (GAAIS-PA). The factor covariance was 0.492, 95% CI [0.455,0.528], SE = 0.019, $z = 26.215$, $p < 0.001$, and the correlation between the two factors was $r = 0.397$, $p < 0.001$. In addition, correlation and regression analyses were conducted between TRI and GAAIS in Study 1. In Study 2, correlation and regression analyses were conducted in a sample group of 300 people with the scores obtained from the 30-item Big Five Inventory-2 Short Form (Soto & John, 2017) consisting of Extraversion, Agreeableness, Conscientiousness, Negative Emotionality, and Open-Mindedness dimensions, the 13-item Corporate Distrust Scale (Adams et al., 2010), and the 6-item General Trust Scale (Yamagishi & Yamagishi). They also found $\alpha = 0.85$ for the positive attitude dimension and $\alpha = 0.82$ for the negative attitude dimension. As a result, it was concluded that the GAAIS performed valid and reliable measurements in this study as well as in the previous scale development study.

Additionally, no studies on the meta-analytic reliability of the GAAIS have been found in the literature. Accordingly, this study aims to generalize the meta-analytic reliability of the GAAIS, which is known to perform valid and reliable measurements and is frequently used in the literature, aims to measure one of today's popular topics, and is one of the first examples developed in the field. Thus, it is thought that this scale will provide preliminary information and shed light on how the reliability of this scale will change according to the variables (different cultures, language, age, rate of females, etc.).

## 2. METHOD

This study follows the meta-analytic method of Vacha-Haase (1998) and aims to generalize the reliability of the General Attitudes Towards Artificial Intelligence Scale. Accordingly, Cronbach's α values for the overall scale and Cronbach's α values for the positive and negative sub-dimensions were generalized. In addition, McDonald's ω coefficients for positive and negative sub-dimensions were also generalized.

### 2.1. Data Collection and Coding Process

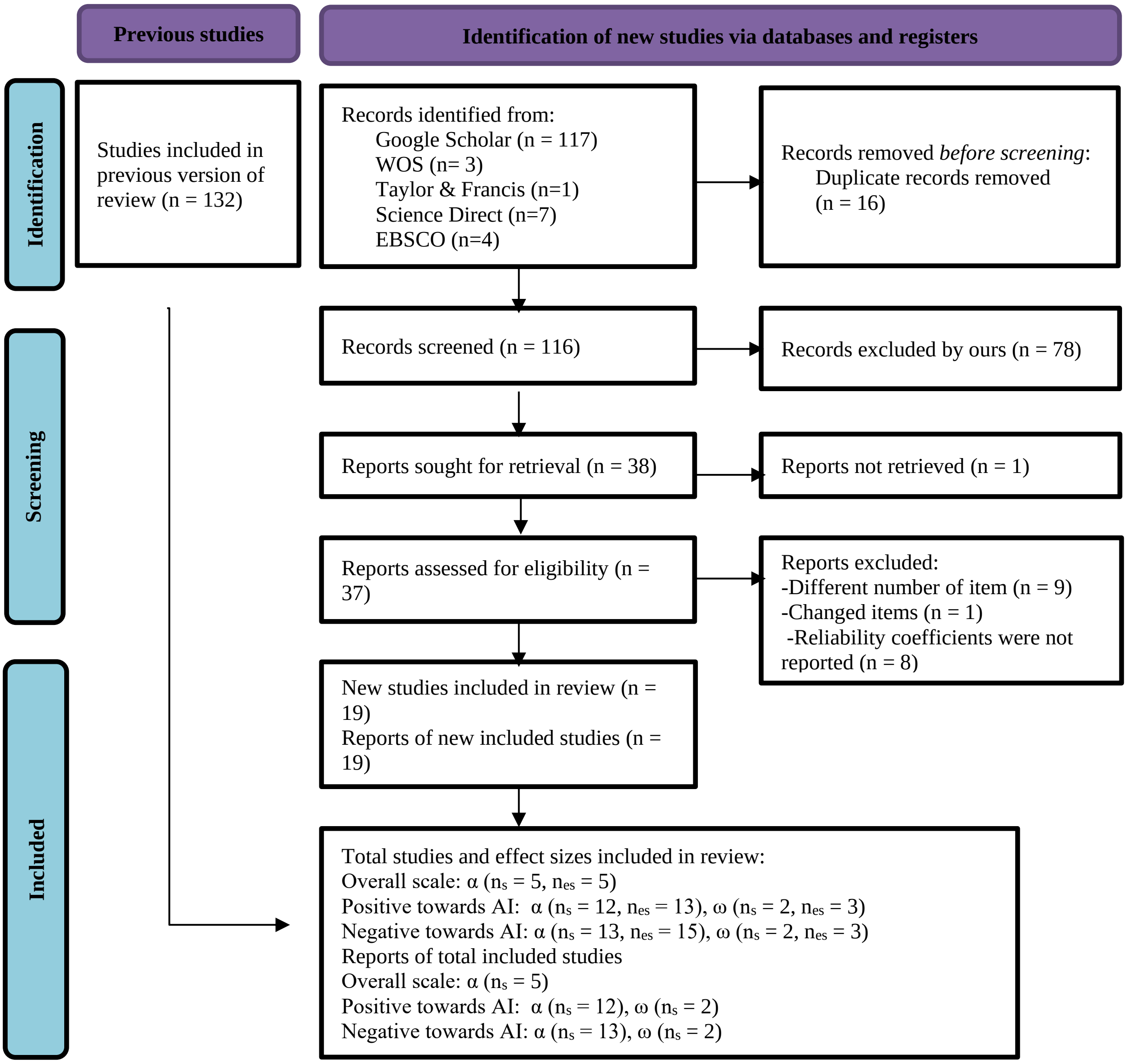
Google Scholar, WOS, Taylor & Francis, Science Direct, and EBSCO databases were searched with the keyword "Attitudes Towards Artificial Intelligence Scale" to access the studies in which the scale was used to perform meta-analytic reliability generalization. As a result of the searches, the full texts of all studies in the databases were examined and included in the meta-analytic reliability generalization if they met the specified criteria. The inclusion criteria can be listed as follows:

- The reporting language of the study can be any language.
- One of the original or adapted forms of the General Attitudes Towards Artificial Intelligence Scale (overall or subscales) must have been used.
- The overall and/or sub-dimensional reliability (Cronbach's α and/or McDonald's ω) of the scale must have been reported.
- The size of the sample to which the scale was applied must have been reported.

The PRISMA 2020 Statement flowchart for the identification, searching, and inclusion of full-text articles reviewed according to these criteria is given in Figure 1 (Page et al., 2021).

According to the PRISMA flowchart shown in Figure 1, 132 studies were identified by searching the determined databases. Of these 132 studies, 16 were eliminated because they were duplicates, 78 were eliminated because they did not use the scale and were related to the subject, and one was eliminated because it could not be accessed. When the remaining 37 studies were examined using the inclusion criteria, it was seen that different numbers of items were used in nine studies (6, 8, 9, 15, 16, and 32 items) and the items were changed in one study. The reliability coefficient was not reported in 8 of the remaining 27 studies. Thus, 19 studies were included in the analysis. In 5 of these 19 studies, 5 Cronbach's α coefficients for the overall scale were generalized. 15 Cronbach's α coefficients from 13 studies for negative attitudes toward AI and 13 Cronbach's α coefficients from 12 studies for positive attitudes towards AI were included in the analysis. Besides, for the other coefficient included in the study, McDonald's ω, three coefficients from 2 studies were generalized for positive and negative subscales.

**Figure 1.** *PRISMA flowchart.*




*$n_s$ = number of studies, $n_{es}$ = number of effect sizes

After the search, the studies selected by the inclusion criteria were coded by two coders. In addition to coding reliability coefficients and sample sizes, the two coders coded some descriptive variables to perform a reliability generalization analysis. These variables were (i) publication citation, (ii) year of publication, (iii) language of publication, (iv) type of research, (v) overall reliability type, (vi) sub-dimension reliability type, (vii) the number of response categories, (viii) scale language, (ix) country, (x) mean age, (xi) standard deviation of age, (xii) study group, (xiii) rate of females, and (xiv) study field. The frequencies of the sub-categories for the categorical variables are given in Table 1.

Since all the research data were coded by two coders, the percentage of inter-coder agreement was calculated according to Miles and Huberman (1994), and the agreement was found to be 100%. After the agreement of the coded data was determined, the data were analyzed.

**Table 1.** *Frequency of studies and effect sizes for Cronbach's α.*

| | | GAAIS-Overall | | GAAIS-Negative | | GAAIS-Positive | |
|---|---|---|---|---|---|---|---|
| | Categories | $n_s$ | $n_{es}$ | $n_s$ | $n_{es}$ | $n_s$ | $n_{es}$ |
| Publish Type | Manuscript | 4 | 4 | 13 | 15 | 12 | 13 |
| | Proceeding | 1 | 1 | - | - | - | - |
| Publish Language | English | 5 | 5 | 12 | 14 | 11 | 12 |
| | Korean | - | - | 1 | 1 | 1 | 1 |
| Scale Language | English | 4 | 4 | 9 | 11 | 8 | 9 |
| | Korean | - | - | 2 | 2 | 2 | 2 |
| | German | - | - | 1 | 1 | 1 | 1 |
| | Turkish | - | - | 1 | 1 | 1 | 1 |
| | Arabic | 1 | 1 | - | - | - | - |
| Likert Type | 3-point | 1 | 1 | - | - | - | - |
| | 5-point | 3 | 3 | 13 | 15 | 12 | 13 |
| | 7-point | 1 | 1 | - | - | - | - |
| Region | Asia | 2 | 2 | 5 | 5 | 5 | 5 |
| | Europe | 1 | 1 | 5 | 6 | 5 | 6 |
| | America | 1 | 1 | 2 | 2 | 2 | 2 |
| Study Group | Adult | 4 | 4 | 7 | 9 | 7 | 9 |
| | Student | 1 | 1 | 3 | 3 | 3 | 3 |
| | Adult and Student | - | - | 2 | 3 | 1 | 1 |
| Research Type | Correlational | 3 | 3 | 8 | 9 | 7 | 7 |
| | Scale Adaptation/Development | - | - | 6 | 6 | 6 | 6 |
| | Descriptive | 1 | 1 | - | - | - | - |
| | Experimental | 1 | 1 | - | - | - | - |
| Study Field | Psychology | 1 | 1 | 7 | 8 | 7 | 8 |
| | Health Science | 1 | 1 | 4 | 4 | 4 | 4 |
| | Management/Communication | 3 | 3 | 2 | 3 | 1 | 1 |

GAAIS-Overall = General Attitude of Artificial Intelligence Scale, GAAIS-Positive = GAAIS Positive Subscale, GAAIS-Negative = GAAIS Negative Subscale

## 2.2. Data Analysis

In the studies handled within the scope of the research, meta-analytic reliability generalization regarding the reliability coefficients for both the overall scale and the subscales was carried out. Meta-analytic reliability generalization, which is an extension of the validity generalization, is used to determine the mean measurement error variance between studies and the sources of the variance (Vacha-Haase, 1998). Meta-analytic reliability generalization analysis was performed with the CMA v.2 program. The reliability coefficient, which is the correlation coefficient, is not suitable for meta-analysis, because the variance depends on correlation. Therefore, it can be combined by transforming and then transforming to a reliability coefficient again (Borenstein et al., 2009; Thompson & Vacha-Haase, 2000). The reliability coefficients obtained in the study were transformed into Fisher's z statistics before being included in the analysis. This transformation method has been suggested in the literature and is often used by meta-analysts (Beretvas et al., 2002). Heterogeneity was examined to determine the type of model to be used in the analyses (Borenstein et al., 2009). Q statistic and its significance (Cochran, 1954), $I^2$ statistic (Higgins & Thomson, 2002), and $\tau^2$ values were analyzed to examine the heterogeneity of the distributions of the studies. The $\tau^2$ estimates were made following the Der-

Simonian Laird (1986) method. Then, publication bias, a crucial issue in meta-analysis studies, was examined. The research also paid attention to scanning in different databases for publication bias. In addition, Rosenthal's (1979) fail and safe method, Begg and Mazumdar's (1994) rank correlation test, Egger's linear regression test (Egger et al., 1997), and Duval and Tweedie's trim and fill method based on funnel plot were used to examine publication bias in the data obtained.

In the study, the α coefficient for the overall reliability of the GAAIS and the reliability generalization of the α and ω coefficients for the subscales were analyzed (Vacha-Haase, 1998). These analyses were carried out according to the random effect model since heterogeneity exists statistically and theoretically (Borenstein et al., 2009). The reliability coefficient may vary depending on the applied group (Crocker & Algina, 1986). Heterogeneity in social sciences is a theoretically expected situation. Because the measures were obtained from individuals living in different regions, speaking different languages, and of different ages and characteristics, within the scope of the research, moderator analyses were performed according to the mixed effect model based on both categorical variables and continuous variables. In the selection of variables, situations where the reliability value may differ in the literature were determined (Aslan et al., 2022; Hess et al., 2014; Lopez-Pina et al., 2015; Ozdemir et al., 2020; Yin & Fan, 2000). Analog ANOVA analysis was performed for each subgroup based on region, study group, research type, and study field variables in Table 1. At this stage, the statistical significance of the reliability coefficients obtained for each subgroup was analyzed. Analog ANOVA is performed to test the significance of the difference in the dependent variable in the subcategories of the categorical independent variable. If there is heterogeneity, this variability may be due to subgroups, so the sources of heterogeneity can be determined by performing Analog to ANOVA. Also, meta-regression analyses were performed for continuous variables such as mean age, standard deviation of age, and rate of females (Caruso & Edwards, 2001; Hess et al., 2014; Youngstrom & Green, 2003). Thus, sample characteristics that could reveal differences in the homogeneity of the group were considered moderator variables in the study (Henson & Thompson, 2002). The significance of the models and explained variance values were reported (Hedges & Pigott, 2004).

## 3. RESULTS

Within the scope of the research, meta-analytic reliability generalization of the reliability coefficients for GAAIS, GAAIS-Negative, and GAAIS-Positive was conducted. For Cronbach's α, the overall scale, and its subscales were considered, while for McDonald's ω, only the subscales were considered because studies reporting McDonald's ω did not calculate this coefficient for the overall scale. For reliability generalizations, publication bias results were first examined and are presented in Table 2.

**Table 2.** *Publication bias.*

| | Overall GAAIS (α) | GAAIS-Negative (α) | GAAIS-Positive (α) | GAAIS-Negative (ω) | GAAIS-Positive (ω) |
|---|---|---|---|---|---|
| Rosenthal Fail-safe | 1782 | 15904 | 13545 | 5038 | 7215 |
| Kendall’s τ | 0.000 | -0.210 | 0.115 | 0.333 | 0.667 |
| Intercept | -2.188 | -2.218 | -3.579 | 8.628 | 7.074 |
| Adjusted studies | 0 | 0 | 0 | 0 | 0 |

When the failsafe-N results regarding publication bias given in Table 2 were examined, it was an indication that publication bias did not exist since the number of missing studies that should be added for the overall reliability coefficient to be non-significant was higher than the criterion value (5k+10) for the overall scale and subscales (Rosenthal, 1979). *k* is the number of studies used in calculating this criterion value. Kendall's τ and Egger regression intercept values were not significant. These tests showed that there was no evidence of funnel plot asymmetry (Begg & Mazumdar, 1994; Egger et al., 1997; Rothstein et al., 2005). Finally, when the results of Duval and Tweedie's trim and fill method were analyzed, it was observed that the number of adjusted studies for the funnel plot to be symmetric was 0 in all results. Accordingly, it could be said that there was no publication bias according to the method by Duval and Tweedie. When all the evidence was analyzed together, it was concluded that there was no publication bias for all coefficients in the overall scale and subscales. In addition, the induction rate was calculated in the study and this rate was found to be 29.63% ((8/27) ×100)) (Vacha-Haase et al., 2000; Sanchez-Meca et al., 2021). The pooled Cronbach's α and McDonald's ω values for the overall scale and subscales and also heterogeneity statistics are presented in Table 3.

**Table 3.** *Results for overall effect sizes and heterogeneity.*

| Overall Coefficients | | | | | | |
|---|---|---|---|---|---|---|
| Coefficients | | *k* | RC [$LL_{RC}$- $UL_{RC}$] | Q | $I^2$ | $\tau^2$ |
| Cronbach's α | Overall | 5 | 0.881* [0.849-0.907] | 12.002* | 66.671 | 0.014 |
| | Negative | 15 | 0.828* [0.807-0.846] | 65.502* | 78.627 | 0.011 |
| | Positive | 13 | 0.863* [0.840-0.883] | 90.917* | 86.801 | 0.020 |
| Mc Donald's ω | Negative | 3 | 0.873* [0.859-0.886] | 6.074* | 67.075 | 0.002 |
| | Positive | 3 | 0.923* [0.916-0.929] | 3.068 | 34.820 | 0.000 |

*$p$<0.05, RC: Reliability Coefficient, $LL_{RC}$: Lower Limit, $UL_{RC}$: Upper Limit, *k:* Number of studies

When analyzing the significance of the reliability values for the total scale and the subscales in Table 3, it was found that all coefficients obtained in both types of reliability coefficients were statistically significant. When Cronbach's α was analyzed, the highest value was obtained in the overall scale, while the lowest value was obtained in the negative subscale. When McDonald's ω values were analyzed, the highest value was obtained in the positive subscale. When both reliability types were analyzed for the subscales, it was observed that McDonald's ω reliability values were higher than Cronbach's α values.

In the heterogeneity values given in Table 3, the Q value was found to be significant for all scales where Cronbach's α was generalized, while for the ω coefficient, it was found to be significant for GAAIS-Negative and not significant for GAAIS-Positive. For $I^2$, another evidence of heterogeneity, GAAIS-overall α, and GAAIS-Negative ω could be considered as moderate heterogeneity indicators. For GAAIS-Negative α and GAAIS-Positive α, $I^2$ could be said to be a high-level heterogeneity indicator. For GAAIS-Positive ω, a low level of heterogeneity was determined (Higgins et al., 2003). When the variance between studies ($\tau^2$) was analyzed, it was seen that all of them except GAAIS-Positive ω were different from 0 and there was a variance between studies. For the ω coefficient of the GAAIS-Positive subscale, it could be said that there was no variance between the studies. In general, heterogeneity existed for both α coefficients and ω coefficients. Forest plots for Cronbach's α coefficient are given for the negative subscale and positive subscale in Appendix 1 and Appendix 2. When the forest plots were examined, it was seen that the Cronbach's alpha (standard error) of the primary studies for both the negative and positive subscales were distributed heterogeneously. The results of moderator analysis with categorical and continuous variables are presented in Table 4.

**Table 4.** *Results for categorical/continuous moderator analysis.*

| GAAIS-Negative | | | | | |
|---|---|---|---|---|---|
| Categorical Moderator | Categories | *k* | α [LLα-ULα] | Q (*df*) | *p* |
| Region | Asia | 5 | 0.811 [0.763-0.850] | | |
| | Europe | 6 | 0.827 [0.788-0.859] | 1.101(2) | 0.577 |
| | America | 2 | 0.848 [0.785-0.894] | | |
| Study Group | Adult | 9 | 0.826 [0.799-0.850] | | |
| | Student | 3 | 0.813 [0.756-0.858] | 0.923(2) | 0.630 |
| | Adult and Student | 3 | 0.843 [0.800-0.878] | | |
| Research Type | Correlational | 9 | 0.831 [0.803-0.855] | 0.127(1) | 0.722 |
| | Scale Devel. /Adapt. | 6 | 0.823 [0.788-0.853] | | |
| Study Field | Psychology | 8 | 0.834 [0.808-0.857] | | |
| | Health Science | 4 | 0.800 [0.752-0.839] | 2.449(2) | 0.294 |
| | Communication | 3 | 0.840 [0.799-0.873] | | |
| | | *k* | β [SE] | $Q_M$ | $Q_R$ |
| Continuous Moderator | Mean Age | 14 | 0.005 [0.002] | 10.098* | 46.870* |
| | Standard Deviation of Age | 14 | 0.007 [0.003] | 4.705* | 52.263* |
| | Rate of Female | 15 | -0.419 [0.105] | 15.795* | 49.707* |
| GAAIS-Positive | | | | | |
| | | *k* | α [LLα-ULα] | Q(*df*) | *p* |
| Region | Asia | 5 | 0.847 [0.814-0.876] | | |
| | Europe | 6 | 0.878 [0.855-0.898] | 3.010(2) | 0.222 |
| | America | 2 | 0.852 [0.800-0.892] | | |
| Study Group | Adult | 9 | 0.870 [0.844-0.892] | 0.253(1) | 0.615 |
| | Student | 3 | 0.857 [0.802-0.897] | | |
| Research Type | Correlational | 7 | 0.852 [0.823-0.877] | 1.638(1) | 0.201 |
| | Scale Devel./Adapt. | 6 | 0.875 [0.849-0.897] | | |
| Study Field | Psychology | 8 | 0.872 [0.846-0.893] | 0.529(1) | 0.467 |
| | Health Science | 4 | 0.855 [0.811-0.889] | | |
| | | *k* | β [SE] | $Q_M$ | $Q_R$ |
| Continuous Moderator | Mean Age | 12 | 0.008 [0.002] | 25.393* | 63.679* |
| | Standard Deviation of Age | 12 | 0.014 [0.004] | 15.857* | 73.216* |
| | Rate of Female | 13 | -0.411 [0.106] | 15,112* | 75.805* |

*$p<0.05$, $LL_\alpha$: Lower Limit, $UL_\alpha$: Upper Limit, *k:* Number of studies, β: Slope, $Q_M$: Q values for model, $Q_E$: Q values for residual

The moderator analysis handled the categorical variables (region, study group, research type, and study field). When analog ANOVA results for the negative subscale and the positive subscale were examined, it was observed that Cronbach's α did not differ significantly in the sub-categories of the variables. In the negative sub-dimension, it was observed that the Cronbach α value obtained in the region-based analyses was the highest in the American region and the lowest in the Asia region. In the analysis based on the study group, the highest reliability value was obtained in the "adult and student" subgroup and the lowest in the student subgroup. In the analysis based on research type, higher reliability values were found in correlational studies. In the study field, the highest reliability value was obtained in communication and the lowest in health science.

In the positive subscale, the highest reliability value was obtained in Europe and the lowest in Asia as a result of the region-based analog to ANOVA. However, this difference was not significant. There was no significant difference between Cronbach's α results based on study groups, but a higher α coefficient was obtained in the analyses conducted with adults. In the research type, the reliability value obtained from correlational studies was relatively higher and the difference was not significant. Finally, it was determined that the reliability value obtained in the field of psychology (based on the field of study) was higher and not statistically significant.

In the moderator analysis, the mean age, standard deviation of age, and rate of females were considered continuous variables. In the negative subscale, the model based on mean age was found to be statistically significant. There was a positive relationship between the mean age and the α coefficient. When the extent to which the mean age explained the variability in the α coefficient was examined by (QM/(OM+QE)) × 100 (Borenstein et al., 2009; Card, 2012), it was seen that the variance explained was 17.73%. It was concluded that the model established by considering the standard deviation of age as a continuous variable was also statistically significant. There was a positive relationship between the standard deviation of age and α coefficient. The variability of the standard deviation of age in α coefficient was 8.26%. In the moderator analysis based on the rate of females, the model was significant and there was a negative relationship between the model and the α coefficient, and the α coefficient decreases as the rate of female participants increases. The rate of the female variable explained 24.11% of the variability in the α coefficient.

When the results of continuous moderator analysis for positive attitudes toward AI were examined, it was seen that mean age, standard deviation of age, and rate of female variables all significantly predicted Cronbach's α. Among these variables, mean age and standard deviation of age positively predicted Cronbach's α, while the rate of females predicted it negatively. It could be said that the α coefficient increased with the rise in mean age and standard deviation of age, and the α coefficient decreased with the increase in the rate of females. When the extent to which mean age explained the variability in the α coefficient was analyzed, it was found that the variance explained was 28.508%. The variance explained by the standard deviation of age was 17.802%. Finally, the variance explained by the rate of females was 16.621%. In both positive and negative attitudes towards AI subscales, it was observed that mean age explained the most variability in the α coefficient.

## 4. DISCUSSION and CONCLUSION

This study aimed to generalize the reliability of the overall GAAIS scale and its subscales. Cronbach's α coefficient was examined for the overall scale, and Cronbach's α and McDonald's ω reliability coefficients were examined for the subscales. When the reliability coefficients were reviewed, it was seen that Cronbach's α coefficient was mainly examined in primary studies. Cronbach's α coefficient is a reliability coefficient that is frequently calculated in the literature (Osburn, 2002; Warrens, 2014). In addition, all reliability coefficients estimated for the overall scale and subscales are above .70 (Nunnally, 1978). Accordingly, it can be said that the overall reliability of the scale is high. McDonald's ω coefficients calculated for the subscales were higher than Cronbach's α. In general, the α coefficient is also lower than the other coefficients. In other words, Cronbach's α is defined as the lower limit of reliability (Kristoff, 1974; Novick & Lewis, 1967).

In our study, moderator analyses were conducted by selecting variables that were frequently examined in the literature in reliability generalization studies. In the region-based analyses, it was determined that the reliability value of the GAAIS scale did not change significantly in the studies conducted in Asian, European, and American regions. However, the reliability values obtained varied according to the regions. This difference was determined as 0.037 in the negative subscale and 0.031 in the positive subscale, but it was not significant. Based on this, it can be stated that the error rates of the responses of people living in different regions to the scale were also different. Obtaining different results in different regions could also be explained by the differentiation in terms of the homogeneity of the distribution of individuals' views on AI practices. In this study, it was observed that the overall reliability values of Europe and America were higher than those of Asia in both the negative subscale and the positive subscale. Similar to the results of this study, there are reliability generalization studies in the literature that calculate lower overall reliability coefficients in Asia (Alcorer-Bruno et al., 2020; Vassar, 2008)

In the categorical moderator analysis based on the study group, reliability estimates were calculated in both subscales in different subgroups, and it was concluded that the difference was not significant. Other studies in the literature conclude that there is no significant difference between the study groups (Thompson & Cook, 2002; Wallace & Weller, 2002). The fact that the lowest reliability value for both subscales was obtained in the student group can be explained by the fact that the students' responses to the scale were more inconsistent or that their views on the AI application were more homogeneous compared to the other group. In the positive subscale, a higher reliability value was obtained in the category of both adults and students. It can be stated that the AI applications that the students encounter in their educational life are also similar compared to those of the other adult groups. This result is related to the heterogeneity of the group and can be explained by the higher value of Cronbach's α. When the reliability for the group of students is generalized, there are also studies in the literature where lower reliability values were obtained compared to more heterogeneous adult or adult-student groups (Eser & Dogan, 2023; Yoruk & Sen, 2022).

Another variable type handled in the study was research type. The reliability values obtained also differed whether the research type was correlational or scale development/adaptation. Although this difference was higher, especially in the positive subscale, it was insignificant in both subscales. In the moderator analysis based on the study field, different coefficients were obtained in different study areas where the research was conducted and this difference was not found to be significant, which is similar to the studies in the literature (Ozdemir et al., 2020). However, in both subscales, the primary studies were conducted mainly in the field of psychology. In the negative and positive attitude subscales, Cronbach's α reliability value obtained from the studies conducted in psychology was higher compared to health science, which may be due to the higher number of studies in the field of psychology. As the number of studies increases, the heterogeneity of the sample may increase. In addition, this result can also be explained by the fact that the groups studying in the field of health are more homogeneous. The characteristics of the sample groups selected in psychology research (occupational status, age groups, family status, education levels, etc.) may differ. In the negative subscale, unlike the positive subscale, data were also obtained in the field of communication and the highest overall reliability value was obtained in this category.

The change in the reliability values of the negative and positive subscales of AI according to the predictor variable, mean age, was analyzed by meta-regression. Mean age was positively correlated with Cronbach's α in the subscales and significantly predicted it. It can be stated that as the average age of the participants increases, their answers are more consistent. A similar relationship exists between the standard deviation of age and Cronbach's α. This is expected because the change in the standard deviation of age indicates that the sample group is heterogeneous in terms of age. As a result of this heterogeneity, it is expected that the overall reliability values will be high. In the literature, it has been observed that there are studies with similar results (Caruso & Edward, 2001; Youngstrom & Green, 2003).

An interesting result obtained from the research is that there is a negative relationship between the rate of females and the scale's Cronbach's α reliability value. The reliability value obtained increases as the rate of females participating in the study decreases. It can also be stated that as the rate of men participating in the study increases, the consistency of the answers given regarding the scale increases. The fact that attitudes towards AI may differ according to gender may also cause this result. Similar to the results of this study, some studies in the literature have found that reliability decreases as the proportion of females increases (Beretvas et al., 2008; Eser & Dogan, 2023). In contrast to the results of this study, Beretvas et al. (2002) determined that reliability decreases as the proportion of men increases in their reliability generalization study.

In this study conducted within the scope of AI, one of the popular topics today, the most cited GAAIS scale was selected. Due to the increase in the number of studies in this field and the fact that the effect sizes are affected by the reliability of the measurement tools, it is vital to examine the reliability of the measurement tools and to determine the change according to the variables specified. With the increase in the number of related studies, moderator analyses can be performed by considering different variables than the variables addressed in this study.

## Declaration of Conflicting Interests and Ethics

The authors declare no conflict of interest. This research study complies with research publishing ethics. The scientific and legal responsibility for manuscripts published in IJATE belongs to the authors.

## Contribution of Authors

**Melek Gülşah Şahin:** Investigation, Methodology, Screening Primary Studies, Coding Primary Studies, Software, Data Analysis, Resources, and Writing-original draft. **Yıldız Yıldırım:** Investigation, Methodology, Screening Primary Studies, Coding Primary Studies, Visualization, Software, Data Analysis, Resources, Writing-original draft.

## Orcid

Melek Gülşah Şahin https://orcid.org/0000-0001-5139-9777
Yıldız Yıldırım https://orcid.org/0000-0001-8434-5062

## REFERENCES

*: Included in the meta-analysis

Alcocer-Bruno, C., Ferrer-Cascales, R., Rubio-Aparicio, M., & Ruiz-Robledillo, N. (2020). The medical outcome study-HIV health survey: A systematic review and reliability generalization meta-analysis. *Research in Nursing & Health*, *43*(6), 610-620. https://doi.org/10.1002/nur.22070

Arslan, K. (2020). Eğitimde yapay zekâ ve uygulamaları [Artificial intelligence and applications in education]. *The Western Anatolia Journal of Educational Sciences, 11*(1), 71-88. https://dergipark.org.tr/tr/pub/baebd/issue/55426/690058

Aslan, Ö.S., Gocen, S., & Sen, S. (2022). Reliability generalization meta-analysis of mathematics anxiety scale for primary school students. *Journal of Measurement and Evaluation in Education and Psychology*, *13*(2), 117-133. https://doi.org/10.21031/epod.1119308

Begg, C.B., & Mazumdar, M. (1994). Operating characteristics of a rank correlation test for publication bias. *Biometrics, 50*(4), 1088. https://doi.org/10.2307/2533446

*Bellaiche, L., Shahi, R., Turpin, M.H., Ragnhildstveit, A., Sprockett, S., Barr, N., ... & Seli, P. (2023). Humans versus AI: Whether and why we prefer human-created compared to AI-created artwork. *Cognitive Research: Principles and Implications, 8*(1), 1-22. https://doi.org/10.1186/s41235-023-00499-6

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

Beretvas, S.N., Suizzo, M.A., Durham, J.A., & Yarnell, L.M. (2008). A reliability generalization study of scores on Rotter's and Nowicki-Strickland's locus of control scales. *Educational and Psychological Measurement*, *68*(1), 97-119. https://doi.org/10.1177/0013164407301529

*Bergdahl, J., Latikka, R., Celuch, M., Savolainen, I., Mantere, E.S., Savela, N., & Oksanen, A. (2023). Self-determination and attitudes toward artificial intelligence: Cross-national and longitudinal perspectives. *Telematics and Informatics*, 82, 102013. https://doi.org/10.1016/j.tele.2023.102013

Borenstein, M., Hedges, L.V., Higgins, J.P., & Rothstein, H.R. (2009). *Introduction to meta-analysis*. John Wiley & Sons.

Breazeal, C. (2004). *Designing sociable robots*. MIT.

Card, N. (2012). *Applied meta-analysis for social science research.* Guilford.

*Carolus, A., Koch, M., Straka, S., Latoschik, M.E., & Wienrich, C. (2023). MAILS-Meta AI Literacy Scale: Development and testing of an AI Literacy Questionnaire based on well-founded competency models and psychological change-and meta-competencies. arXiv preprint. https://doi.org/10.48550/arXiv.2302.09319

Caruso, J.C., & Edwards, S. (2001). Reliability generalization of the Junior Eysenck Personality Questionnaire. *Personality and Individual Differences,* 31, 173-184. https://doi.org/10.1016/S0191-8869(00)00126-4

Cochran, W.G. (1954). The combination of estimates from different experiments. *Biometrics,* 10, 101–129. https://doi.org/10.2307/3001666

*Cruz, J.P., Sembekova, A., Omirzakova, D., Bolla, S.R., & Balay-odao, E.M. (2023). General attitudes towards and readiness for medical artificial intelligence among medical and health sciences students in Kazakhstan. https://doi.org/10.2196/preprints.49536.

*Darda, K., Carre, M., & Cross, E. (2023). Value attributed to text-based archives generated by artificial intelligence. *Royal Society Open Science*, 10: 220915. https://doi.org/10.1098/rsos.220915

DerSimonian, R., & Laird, N. (1986). Meta-analysis in clinical trials. *Controlled clinical trials, 7*(3), 177-188. https://www.biostat.jhsph.edu/~fdominic/teaching/bio656/references/sdarticle.pdf

Eser, M.T., & Dogan, N. (2023). Life Satisfaction Scale: A meta-analytic reliability generalization study in Turkey sample. *Turkish Psychological Counseling and Guidance Journal*, *13*(69), 224-239. https://doi.org/10.17066/tpdrd.1223320mn

*Gabbiadini, A., Dimitri, O., Cristina, B., & Anna, M. (2023). Does ChatGPT pose a threat to human identity. *SSRN*, 4377900. https://doi.org/10.2139/ssrn.4377900

*Gozzo, M., Woldendorp, M.K., & De Rooij, A. (2021, December). Creative collaboration with the "brain" of a search engine: Effects on cognitive stimulation and evaluation apprehension. In *International Conference on ArtsIT, Interactivity and Game Creation* (pp. 209-223). Springer International Publishing.

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

*Hadlington, L., Binder, J., Gardner, S., Karanika-Murray, M., & Knight, S. (2023). The use of artificial intelligence in a military context: Development of the Attitudes Toward AI in Defense (AAID) Scale. *Frontiers in Psychology*, 14, 1164810. https://doi.org/10.3389/fpsyg.2023.1164810

Hedges, L.V., & Pigott, T.D. (2004). The power of statistical tests for moderators in meta-analysis. *Psychological Methods, 9*(4), 426-445. https://doi.org/10.1037/1082-989x.9.4.426

*Heim, S., & Chan-Olmsted, S. (2023). Consumer trust in AI–human news collaborative continuum: preferences and influencing factors by news production phases. *Journalism and Media*, *4*(3), 946-965. https://doi.org/10.3390/journalmedia4030061

Henson, R.K., & Thompson, B. (2002). Characterizing measurement error in scores across studies: Some recommendations for conducting "reliability generalization" studies.

*Measurement and Evaluation in Counseling and Development*, *35*(2), 113-127. https://doi.org/10.1080/07481756.2002.12069054

Hess, T.J., McNab, A.L., & Basoglu, K.S. (2014). Reliability generalization of perceived ease of use, perceived usefulness, and behavioral intentions. *MIS Quarterly*, 38, 1-28. https://doi.org/10.25300/MISQ/2014/38.1.01

Higgins, J.P.T., & Thompson, S.G. (2002), Quantifying heterogeneity in a meta-analysis. *Statistics in Medicine*, 21, 1539-1558. https://doi.org/10.1002/sim.1186

Hopcan, S., Turkmen, G., & Polat, E. (2023). Exploring the artificial intelligence anxiety and machine learning attitudes of teacher candidates. *Education and Information Technologies*, 1-21. https://doi.org/10.1007/s10639-023-12086-9

Huang, S.P. (2018). Effects of using artificial intelligence teaching system for environmental education on environmental knowledge and attitude. *Eurasia Journal of Mathematics, Science and Technology Education*, *14*(7), 3277-3284. https://doi.org/10.29333/ejmste/91248

Kandlhofer, M., Steinbauer, G., Hirschmugl-Gaisch, S., & Huber, P. (2016, October). Artificial intelligence and computer science in education: From kindergarten to university. In *2016 Institute of electrical and electronics engineers - Frontiers in education conference (IEEE-FIE)* (pp. 1-9). IEEE. https://doi.org/10.1109/FIE.2016.7757570

Kaplan, A., & Haenlein, M. (2019). Siri, Siri, in my hand: Who's the fairest in the land? On the interpretations, illustrations, and implications of artificial intelligence. *Business Horizons*, *62*(1), 15-25. https://doi.org/10.1016/j.bushor.2018.08.004

*Kaya, F., Aydin, F., Schepman, A., Rodway, P., Yetisensoy, O., & Demir Kaya, M. (2022). The roles of personality traits, AI anxiety, and demographic factors in attitudes toward artificial intelligence. *International Journal of Human–Computer Interaction*, 1-18. https://doi.org/10.1080/10447318.2022.2151730

Kieslich, K., Lünich, M., & Marcinkowski, F. (2021). The threats of artificial intelligence scale (TAI) development, measurement and test over three application domains. *International Journal of Social Robotics*, 13, 1563-1577. https://doi.org/10.1007/s12369-020-00734

Kristof, W. (1974). Estimation of reliability and true score variance from a split of a test into three arbitrary parts. *Psychometrika*, 39, 491-499. https://doi.org/10.1007/BF02291670

*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), 1-8. https://doi.org/10.1186/s12912-022-01048-0

*Kwak, Y., Seo, Y.H., & Ahn, J.W. (2022). Nursing students' intent to use AI-based healthcare technology: Path analysis using the unified theory of acceptance and use of technology. *Nurse Education Today*, 119: 105541. https://doi.org/10.1016/j.nedt.2022.105541

McCarthy, J. (2004). *What is artificial intelligence?*. http://www.formal.stanford.edu/jmc/whatisai/

*Mohamed, H.A., Awad, S.G., Eldiasty, N.E.M.M, & ELsabahy, H.E. (2023). Effect of the artificial intelligence enhancement program on head nurses' managerial competencies and flourishing at work. *Egyptian Journal of Health Care, 14*(1), 624-645. https://doi.org/10.21608/EJHC.2023.287188

Nica, E., Sabie, O.M., Mascu, S., & Luţan, A.G. (2022). Artificial intelligence decision-making in shopping patterns: Consumer values, cognition, and attitudes. *Economics, Management and Financial Markets*, *17*(1), 31-43. https://doi.org/10.22381/emfm17120222.

*Nguyen, E. (2023). Trust and algorithmic decision making. *UC Santa Barbara*, *3*(2022), 1-15. https://escholarship.org/content/qt5z86t0dx/qt5z86t0dx.pdf

Novick, M.R., & Lewis, C.L. (1967). Coefficient alpha and the reliability of composite measurements. *Psychometrika*, 32, 1-13. https://doi.org/10.1007/BF02289400

Nunnally, J.C. (1978). *Psychometric theory* (2nd ed.). McGraw-Hill.

Osburn, H.G. (2000). Coefficient alpha and related internal consistency reliability coefficients. *Psychological Methods*, *5*(3), 343–355. https://doi.org/10.1037/1082-989X.5.3.343

Ozdemir, V., Yildirim, Y., & Tan, S. (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

Page, M.J., McKenzie, J.E., Bossuyt, P.M., Boutron I., Hoffmann, T.C., Mulrow, C.D., …, & Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ*, 372: 71. https://doi.org/10.1136/bmj.n71

Persson, A., Laaksoharju, M., & Koga, H. (2021). We mostly think alike: Individual differences in attitude towards AI in Sweden and Japan. *The Review of Socionetwork Strategies*, *15*(1), 123-142. https://doi.org/10.1007/s12626-021-00071-y

Pinto dos Santos, D., Giese, D., Brodehl, S., Chon, S.H., Staab, W., Kleinert, R., ... & Baeßler, B. (2019). Medical students' attitude towards artificial intelligence: A multicentre survey. *European radiology*, 29, 1640-1646. https://doi.org/10.1007/s00330-018-5601-1

Polesie, S., Gillstedt, M., Kittler, H., Lallas, A., Tschandl, P., Zalaudek, I., & Paoli, J. (2020). Attitudes towards artificial intelligence within dermatology: An international online survey. *British Journal of Dermatology*, *183*(1), 159-161. https://doi.org/ 10.1111/bjd.18875

Rothstein, H.R., Sutton, A.J., & Borenstein, M. (2005). *Publication bias in meta-analysis: Prevention, assessment and adjustments*. John Wiley & Sons.

Rosenthal, R. (1979). The ‘‘file drawer problem’’ and tolerance for null results. *Psychological Bulletin*, 86, 638–641. https://doi.org/10.1037/0033-2909.86.3.638

*Saddique, F., Usman, M., Nawaz, M., & Mushtaq, N. (2020). Entrepreneurial orientation and human resource management: The mediating role of Artificial Intelligence. *Elementary Education Online*, *19*(4), 4969-4978. https://doi.org/10.17051/ilkonline.2021.05.777

Sánchez-Meca J, Marín-Martínez F, López-López JA, … & López-Nicolás, P. (2021). Improving the reporting quality of reliability generalization meta-analyses: The REGEMA checklist. *Research Synthesis Methods*, 12, 516-536. https://doi.org/10.1002/jrsm.1487

*Schepman, A., & Rodway, P. (2020). Initial validation of the general attitudes towards Artificial Intelligence Scale. *Computers in Human Behavior Reports*, 1, 100014. https://doi.org/10.1016/j.chbr.2020.100014

*Schepman, A., & Rodway, P. (2022). The General Attitudes towards Artificial Intelligence Scale (GAAIS): Confirmatory validation and associations with personality, corporate distrust, and general trust. *International Journal of Human–Computer Interaction*, *39*(13), 2724-2741. https://doi.org/10.1080/10447318.2022.2085400

*Seo, Y.H., & Ahn, J.W. (2022). The validity and reliability of the Korean version of the General Attitudes towards Artificial Intelligence Scale for nursing students. *The Journal of Korean Academic Society of Nursing Education*, *28*(4), 357-367. https://doi.org/10.5977/jkasne.2022.28.4.357

Sindermann, C., Sha, P., Zhou, M., Wernicke, J., Schmitt, H.S., Li, M., ... & Montag, C. (2021). Assessing the attitude towards artificial intelligence: Introduction of a short measure in German, Chinese, and English language. *KI-Künstliche Intelligenz*, 35, 109-118. https://doi.org/10.1007/s13218-020-00689-0

Thompson, B., & Cook, C. (2002). Stability of the reliability of libqual+™ scores a reliability generalization meta-analysis study. *Educational and Psychological Measurement*, *62*(4), 735-743. https://doi.org/10.1177/0013164402062004013

Thompson, B., & Vacha-Haase, T. (2000). Psychometrics is datametrics: The test is not reliable. *Educational and Psychological Measurement*, 60, 174-195. https://doi.org/10.1177/00131640021970448

Turkle, S., Breazeal, C., Dasté, O., & Scassellati, B. (2006). Encounters with kismet and cog: Children respond to relational artifacts. *Digital media: Transformations in human communication*, 120. http://web.mit.edu/people/sturkle/encounterswithkismet.pdf

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

Vassar, M.A. (2008). Note on the score reliability for the Satisfaction with Life Scale: An RG study. *Soc Indic Res*, 86, 47–57. https://doi.org/10.1007/s11205-007-9113-7

Waliszewski, K., & Warchlewska, A. (2020). Attitudes towards artificial intelligence in the area of personal financial planning: A case study of selected countries. *Entrepreneurship and Sustainability Issues, 8*(2), 399-420. https://doi.org/10.9770/jesi.2020.8.2(24)

Wallace, K.A., & Wheeler, A.J. (2002). Reliability generalization of the life satisfaction index. *Educational and Psychological Measurement*, *62*(4), 674-684. https://doi.org/10.1177/0013164402062004009

*Wang, H., Sun, Q., Gu, L., Lai, K., & He, L. (2022). Diversity in people's reluctance to use medical artificial intelligence: Identifying subgroups through latent profile analysis. *Frontiers in Artificial Intelligence*, 5: 1006173. https://doi.org/10.3389/frai.2022.1006173

Warrens, M.J. (2014). On Cronbach's alpha as the mean of all possible-split alphas. *Advances in Statistics*. 742863. https://doi.org/10.1155/2014/742863

Yin, P., & Fan, X. (2000). Assessing the reliability of Beck Depression Inventory scores: Reliability generalization across studies. *Educational and Psychological Measurement, 60*(2), 201-223. https://doi.org/10.1177/00131640021970466

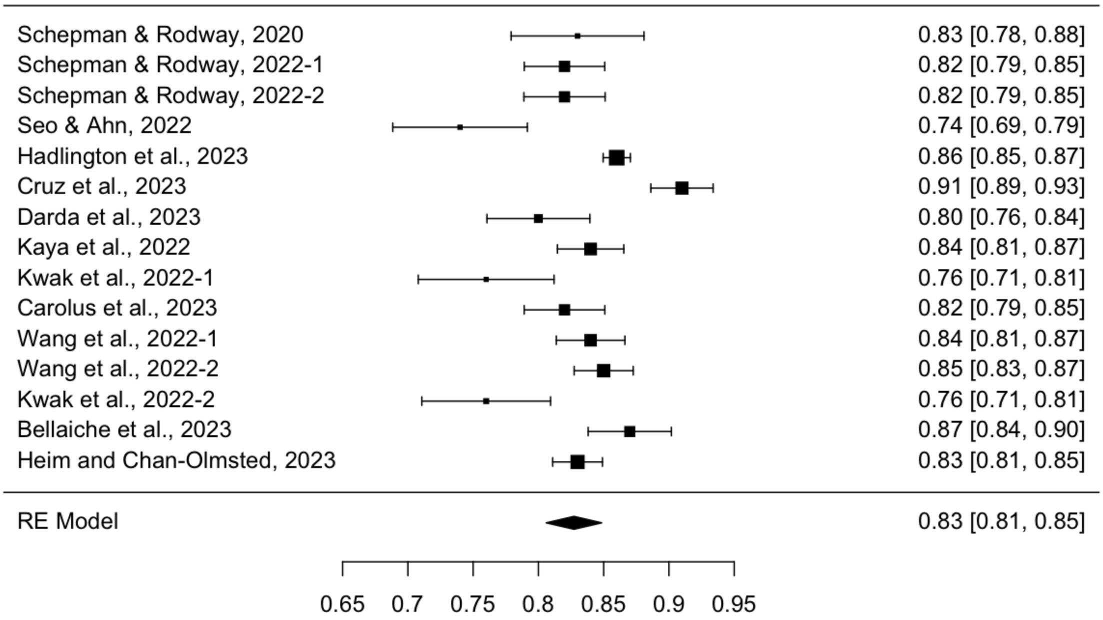
Youngstrom, E.A., & Green, K.W. (2003). Reliability generalization of self-report of emotions when using the Differential Emotions Scale. *Educational and Psychological Measurement*, *63*(2), 279-295. https://doi.org/10.1177/00131644032532

Yoruk, S., & Sen, S. (2023). A reliability generalization meta-analysis of the creative achievement questionnaire. *Creativity Research Journal*, *35*(4), 714-729. https://doi.org/10.1080/10400419.2022.2148073

Yuzbasioglu, E. (2021). Attitudes and perceptions of dental students towards artificial intelligence. *Journal of Dental Education*, *85*(1), 60-68. https://doi.org/10.1002/jdd.12385

# APPENDIX

**Appendix 1.** *Forest plot for negative subscale (Cronbach's α).*




**Appendix 2.** *Forest plot for positive subscale (Cronbach's α).*