Current Psychology
https://doi.org/10.1007/s12144-025-08348-z

Measurement invariance of the artificial intelligence attitude scale (AIAS-4): cross-cultural studies in Poland, the USA, and the UK

Wiesław Talik[1] · Elżbieta Barbara Talik[2] · Simone Grassini[3,4]

Accepted: 30 July 2025

Abstract

This study investigates the measurement invariance of the Artificial Intelligence Attitude Scale (AIAS-4) across three countries: Poland (N=511), the UK (N=230), and the USA (N=300). The AIAS-4 is a concise measure of general attitude towards artificial intelligence developed by Simone Grassini. The validation process consisted of two steps. Firstly, the cultural adaptation of the Polish version of the AIAS-4 was conducted, which was necessary for subsequent cross-cultural comparisons. The factor structure, internal, and construct validity of the Polish AIAS-4 were estimated. Next, multigroup confirmatory factor analysis was conducted to compare measurement invariance across three populations: Polish, British, and American. The results confirm that the factor structure of the Polish version of the scale is identical to the original one, with high reliability and acceptable validity. Multigroup confirmatory factor analysis supports configural invariance and partial metric invariance. However, equivalence at the scalar level has not been achieved in cross-cultural research.

Keywords Attitude towards AI · Polish adaptation of AIAS-4 · Cross-cultural measurement invariance

Introduction

The advancement of Artificial Intelligence (AI) technology has significantly impacted various societal domains, reflecting a sudden evolution in the technological landscape and a rapid integration of technology into daily life (Huang et al., 2022). The fast progression of AI technology has sparked extensive scholarly discourse, particularly regarding its ethical, sociopolitical, and economic implications (Neudert et al., 2020). Among the most pressing concerns is the potential for AI to reduce the need for human workers (Webster & Ivanov, 2019). This aspect is crucial for psychological research, which seeks to understand the impact of such technological advancements on human labor dynamics and the associated psychological stressors. Additionally, the broader implications of AI's integration into society encompass a range of ethical and security risks (Rodrigues, 2020). These include potential violations of privacy, ethical dilemmas in AI decision-making, and the exacerbation of societal disparities through algorithmic biases (Kostick-Quenet et al., 2022). The psychological ramifications of these issues are significant, warranting in-depth exploration to mitigate adverse effects on mental health and societal well-being.

In exploring cross-cultural differences in attitudes towards AI, it is crucial to consider established theoretical frameworks that elucidate how cultural factors shape individual and societal perceptions of technology. Hofstede's cultural dimensions theory (Hofstede, 1984, 2011) provides a robust foundation by outlining key cultural characteristics, such as uncertainty avoidance, power distance, and individualism versus collectivism, that researchers have frequently employed to analyze technology acceptance across cultures (Jan et al., 2024).

✉ Wiesław Talik
talik@kul.pl

Elżbieta Barbara Talik
etalik@kul.pl

Simone Grassini
simone.grassini@uib.no

1 Institute of Social Science, The John Paul II Catholic University of Lublin, Lublin, Poland

2 Institute of Psychology, The John Paul II Catholic University of , Lublin, Poland

3 Department of Psychosocial Science, University of Bergen, Bergen, Norway

4 Cognitive and Behavioral; Neuroscience Lab, University of Stavanger, Stavanger, Norway

Published online: 09 August 2025

Countries with high uncertainty avoidance may exhibit heightened skepticism or cautiousness towards AI technologies, given their preference for predictability and clear guidelines (Hofstede, 2011). Conversely, societies with lower uncertainty avoidance tend to approach technological innovations more openly and optimistically (Kummer & Schmiedel, 2016). Similarly, cultures with lower power distance, emphasizing egalitarian values, might view AI as an empowering instrument for democratizing information and decision-making processes, whereas those with higher power distance may perceive AI as reinforcing existing hierarchical structures and authority (Hofstede, 2011).

Furthermore, research has highlighted that nations differ significantly in cultural characteristics, technological maturity, digital infrastructure, and the intensity of AI-related research and development activities (Vincent-Lancrin & Van der Vlies, 2020). Such variations influence public attitudes and acceptance of AI, as technologically advanced nations typically possess higher exposure to and familiarity with AI applications, potentially resulting in more informed attitudes.

Recognizing these complexities, the current study explores measurement equivalence, examining how attitudes toward AI are conceptualized and operationalized across culturally and technologically diverse contexts. Specifically, this research investigates comparative perceptions in three distinct countries: Poland, the United Kingdom, and the United States of America. By selecting countries with different technological landscapes, cultural profiles, and histories of AI research, the study contributes to a richer understanding of how artistic and technological contexts interact in shaping public attitudes toward AI.

For researchers, particularly those conducting cross-cultural studies, verifying measurement equivalence (or measurement invariance) is crucial. This determines how a method used in one country is equally understood and evaluated in another. Establishing measurement invariance of instruments allows for further testing of research hypotheses on the relations between constructs or mean differences at the initial stage of data analysis (comp. Byrne et al., 1989; Lubiewska & Głogowska, 2018).

Multiple levels of verification exist for measurement invariance (MI). *Configural invariance* refers to the equality of factor structure, specifically the number of factors and pattern of factor loadings, across different groups. This will help determine whether we measure the same or different constructs in these groups. The second level refers to *weak or metric invariance*, which requires equal factor loadings across groups. In other words, this level of measurement equivalence answers whether the unit of measurement in the groups analysed is comparable, as indicated by the equality of the factor loadings in the two groups. The third level is *strong or scalar invariance*. This refers to equal item intercepts across groups. It ensures that individuals with the same level of a trait receive the same score when measured using the scale. *Strict invariance* refers to equal item error variances across group. Achieving context residual invariance in research is challenging for groups, so the final type of MI is rarely tested (Cieciuch & Davidov, 2015; Lubiewska & Głogowska, 2018).

The article aims to evaluate the measurement equivalence of the AIAS-4 scale which will be described in detail below in the *Method section*.

Materials and methods

Sample

This study investigates the measurement invariance of the Artificial Intelligence Attitude Scale (AIAS-4) across three countries: the UK, the USA and Poland. The UK group consisted of 230 adults aged 18–76 years (M=40.2, SD=14.61), including 114 women and 114 men (2 respondents refused to answer), who were recruited from the online platform Prolific. The US group consisted of 300 adults aged 18–81 years (M=41.5, SD=15.05), including 142 women, 151 men (7 respondents refused to answer). The Polish group consisted of 1021 adults aged 18–83 years (M=48.41, SD=17.01). This sample was randomly divided into two groups: the first, N=510, on which EFA and a preliminary assessment of the psychometric validity of the Polish version of the AIAS-4 were carried out (study 1); and the second, N=511, on which CFA was carried out to confirm the factor structure (study 2).

Instrument

To evaluate the measurement equivalence in the UK, the USA and Poland, the AIAS-4 scale was used. The AIAS-4 is a 4-item questionnaire validated by Grassini (2023) and has satisfactory psychometric properties. The AIAS-4 assesses individuals' overall attitudes toward AI and examines their perceptions of how AI impacts their personal lives, professional activities, and society in general. The development of the scale was grounded in established theoretical frameworks widely used to predict technology acceptance and usage behavior. Specifically, the Technology Acceptance Model (TAM; Davis, 1989) informed the inclusion of items focusing on two central determinants: perceived usefulness (the extent to which individuals believe using the technology enhances their performance) and perceived ease of use (how effortless the use of technology is perceived to be). Moreover, the Unified Theory of Acceptance and Use of Technology (UTAUT; Venkatesh et al., 2003), which integrates constructs from multiple theoretical models, provided additional conceptual guidance. The UTAUT framework emphasizes four core predictors of technology acceptance and adoption: performance expectancy (the belief that technology use will lead to improved job performance), effort expectancy (the ease associated with using the technology), social influence (the degree to which individuals perceive essential others encourage technology use), and facilitating conditions (the availability of resources and support necessary for technology usage). Furthermore, recent empirical studies and theoretical investigations addressing the societal implications of AI were considered in the development of the scale (Grassini, 2023).

Procedure

To assess the equivalence of a measure, the first step is to check whether the model has a similar factor structure and is a good fit to the data in all the countries being compared. Grassini (2023) confirmed a one-factor solution in two studies. The first validation study comprised 230 UK adults aged 18–76 years (M=40.2, SD=14.6) who were recruited from the online platform Prolific. The exploratory factor analysis confirmed the one-factor structure of the scale, with factor loadings ranging from 0.78 to 0.89. The internal consistency was good, as indicated by a Cronbach's alpha of $\alpha=0.902$ and McDonald's Omega of $\omega=0.904$. The measure's validity was confirmed through correlation analysis with the attitude components of the Media and Technology Usage and Attitudes Scale (MTUAS). The results indicated a weak to moderate association between the AIAS-4 score and the MTUAS score.

The one-factor structure was supported by CFA in the second study, which included 300 US adults aged 18–76 years (M=40.2, SD=14.6) ($\chi2(2)=2.49$, p=.289; CFI=0.999, TLI=0.998; RMSEA=0.0285, 90% CI [0 0.122]). Similar results for the reliability and validity of the scale were obtained, as in Study 1.

To address whether the factor structure of the AIAS-4 scale remains consistent in a Polish sample, a cultural adaptation of the test was conducted.

The polish version of the AIAS-4

An adaptation of the AIAS-4 was carried out based on the items reported in the original validation article (Grassini, 2023). The development of the Polish version followed cultural adaptation standards (Hornowska & Paluchowski, 2004). The study began with a theoretical analysis of the construct, followed by linguistic and psychometric adaptations. The latter included preliminary adaptation studies using exploratory factor analysis (EFA) for psychometric data analysis (study 1) and proper studies using confirmatory factor analysis (study 2) to confirm the final factor structure of the AIAS-4.

Linguistic adaptation was carried out using the collaborative and iterative translation method. Two psychologists– a PhD academic teacher and a professional translator of English– independently translated the scale into Polish. The best version of the questionnaire was subsequently identified and agreed upon. The back-translation procedure was abandoned due to its shortcomings, mainly the ambiguity of the translated words and expressions and the difficulty of achieving literal translation (cf. Douglas & Craig, 2007; van de Vijver & Hambleton, 1996). The retranslation procedure, although it is traditionally recommended and used (cf. Brislin, 1976; van de Vijver & Hambleton, 1996), has been criticized. Back translation provides a direct or literal translation from one language to another. Nevertheless, it is feasible to translate a text without accurately conveying the intended meaning. Furthermore, the back translation methodology employs an etic approach to linguistic translation, whereby an equivalent word or construct is presumed to exist in the target language.

Consequently, in certain instances, a comprehensive phrase may be indispensable for elucidating a construct. In light of the aforementioned limitations of back translation, we decided to employ a collaborative and iterative translation method, which is currently recommended in cross-cultural research (cf. Douglas & Craig, 2007).The instructions for providing answers and calculating scores remained the same as those for the original measure.

A total of 1021 adults aged 18–83 years (M=48.41, SD=17.01) were interviewed online via the computer-assisted web interview (CAWI) method by the Pollster Research Institute, a research company specializing in scientific research using new technologies. All procedures were following the Declaration of Helsinki. Informed consent was obtained from all individual adult participants included in the study.

To carry out further steps of statistical analysis, the sample was randomly divided into two groups: the first, N=510, on which EFA and a preliminary assessment of the psychometric validity of the Polish version of the AIAS-4 were carried out (study 1); and the second, N=511, on which CFA was carried out to confirm the factor structure (study 2). To ascertain the homogeneity of the two samples, Pearson's chi-squared test coefficients were calculated for gender (χ^2 = 0.71; p=.40), place of residence (χ^2 = 5.14; p=.53) and education (χ^2 = 0.72; p=.87). A Student's t-test for age was also calculated (t=0.10; p=.46). It indicates that there is no significant difference between the two samples with respect to these demographic variables. Therefore, the samples can be considered homogeneous.

Study 1

Sample

The study sample comprised 510 individuals aged between 18 and 83 years (M=48.46, SD=17.34), with 252 women (49.4%) and 258 men (50.6%). The participants had either vocational or high education in equal proportions (27.3%), while only a small proportion had primary education (7.1%). Additionally, 38.4% of the participants had secondary education. A total of 37.6% of the participants came from rural areas, while the others were from the city.

Statistical analysis

Following the procedure used by the author of the AIAS-4, we replicated the consecutive stages of psychometric adaptation of the measure. All calculations were performed in SPSS version 28.

Exploratory factor analysis

An exploratory factor analysis was then carried out using principal component analysis. The choice of the factor analysis model was formally supported by the Kaiser–Meyer–Olkin (KMO) index (0.849) and Bartlett's sphericity test ($\chi 2$=1557.879, p<.001). The analysis revealed a single factor accounting for 81.329% of the variance. The item factor loadings were very good (Table 1).

Internal consistency

In the following step, we evaluated the reliability of the complete questionnaire. The internal consistency coefficient of Cronbach's α was 0.923, and that of McDonald's Omega ω was 0.923, indicating excellent internal consistency.

Additionally, we evaluated the discriminatory power of each item, as shown in Table 2, through item-total correlations.

Table 1 Item–Factor loadings for the one-factor solution using principal component analysis (N=510)

Item EN [PL]	Factor I
1. I believe that AI will improve my life. [Uważam, że sztuczna inteligencja poprawi moje życie.]	0.923
2. I believe that AI will improve my work. [Uważam, że sztuczna inteligencja usprawni moją pracę.]	0.872
3. I think I will use AI technology in the future. [Myślę, że będę korzystać z technologii sztucznej inteligencji w przyszłości.]	0.901
4. I think AI technology is positive for humanity. [Myślę, że technologia sztucznej inteligencji jest pozytywna dla ludzkości/przynosi ludzkości korzyści.]	0.910

Table 2 Reliability and discriminatory power of the AIAS-4 items

	Item–total correlation	Cronbach's α if item deleted
1. I believe that AI will improve my life. [Uważam, że sztuczna inteligencja poprawi moje życie.]	0.856	0.888
2. I believe that AI will improve my work. [Uważam, że sztuczna inteligencja usprawni moją pracę.]	0.777	0.915
3. I think I will use AI technology in the future. [Myślę, że będę korzystać z technologii sztucznej inteligencji w przyszłości.]	0.821	0.900
4. I think AI technology is positive for humanity. [Myślę, że technologia sztucznej inteligencji jest pozytywna dla ludzkości/przynosi ludzkości korzyści.]	0.834	0.895

All the correlations were statistically significant (p<.001) and ranged from moderate to high, indicating good discriminatory power of the scale items.

Convergent validity

We then assessed the validity of the Polish version of the AIAS-4 using Pearson's r correlations with the following measures:

1. The General Attitudes towards Artificial Intelligence Scale (GAAIS), developed by Schepman and Rodway (2023) in 2020, is a self-administered questionnaire used to measure general attitudes toward AI. The scale consists of 20 items rated on a five-point scale and two subscales: the Positive GAAIS (12 items) and the Negative GAAIS (8 items). The Positive GAAIS has a reliability of α=0.90, and the Negative GAAIS has a reliability of α=0.84.
2. Two statements from the survey on knowledge of artificial intelligence (*To what extent do you agree or disagree with the statement: I understand the concept of artificial intelligence?*) and its impact on life (*To what extent do you think artificial intelligence affect your life?*).

All the correlations were significant, positive, and moderate, indicating that attitudes towards AI are more positive when there is greater knowledge of AI (r=.345; p<.01) and its impact on human life (r=.449; p<.01). The moderate strength of the association is likely attributable to the fact that both knowledge of AI and its impact on life were estimated using a single item.

The AIAS-4 score showed a positive correlation with the GAAIS score, indicating a relationship between Positive GAAIS (r=.829; p<.001) and Negative GAAIS (r=.481;

$p<.001$) aspects of artificial intelligence and general attitudes towards AI.

Study 2

Sample

The second part of the sample ($N=511$), aged between 18 and 83 years ($M=48.36$, $SD=16.70$), included 266 women (52.1%) and 245 men (47.9%) and was subjected to CFA. The educational structure of the respondents was similar to that in Study 1, with 29.2% having vocational education, 27.2% having high education, 36.2% having secondary education, and only 7.4% having primary education. 43.6% of the participants came from rural areas, while the others were from the city.

Statistical analysis

We conducted a CFA to confirm the structural factors of the AIAS-4. The calculations were made in IBM SPSS AMOS, version 29. The requirements for confirmatory factor analysis were met, including linearity of the relationship between variables (confirmed by scatter plots) and sample randomization. To determine model fit, we have applied commonly used statistics (Yu & Muthén, 2002).

Results

The results showed that the one-factor results had the best fit to the data (Fig. 1).

Goodness-of-fit indices for the model: $\chi^2 = 0.36$; $p<.55$; $df=1$; $\chi^2/df=0.36$; RMSEA=0.00; GFI=1.00; AGFI=1.00; CFI=1.00; NFI=1.00; PCLOSE=0.74; CN Hoeltera=5383.

The errors of items 2 and 4 have been correlated because analysis of the residuals (modification indices) shows a strong correlation between them. Correlating these errors improved the quality of the model fit to the data.

Factor loadings are high (from 0.86 to 0.91). Additionally, the percentage of explained variance (multiple correlation coefficient R2) exceeded 0.5. The Polish version of the AIAS-4 has excellent reliability, with a McDonald's Omega reliability coefficient of $\omega=0.931$ and a Cronbach's alpha of $\alpha=0.931$.

After confirming that the factor structure of the Polish version of the AIAS is identical to that of the British and American populations, we proceeded to assess measurement equivalence.

Measurement invariance

The cross-cultural measurement equivalence of the AIAS was assessed by comparing results obtained from Polish

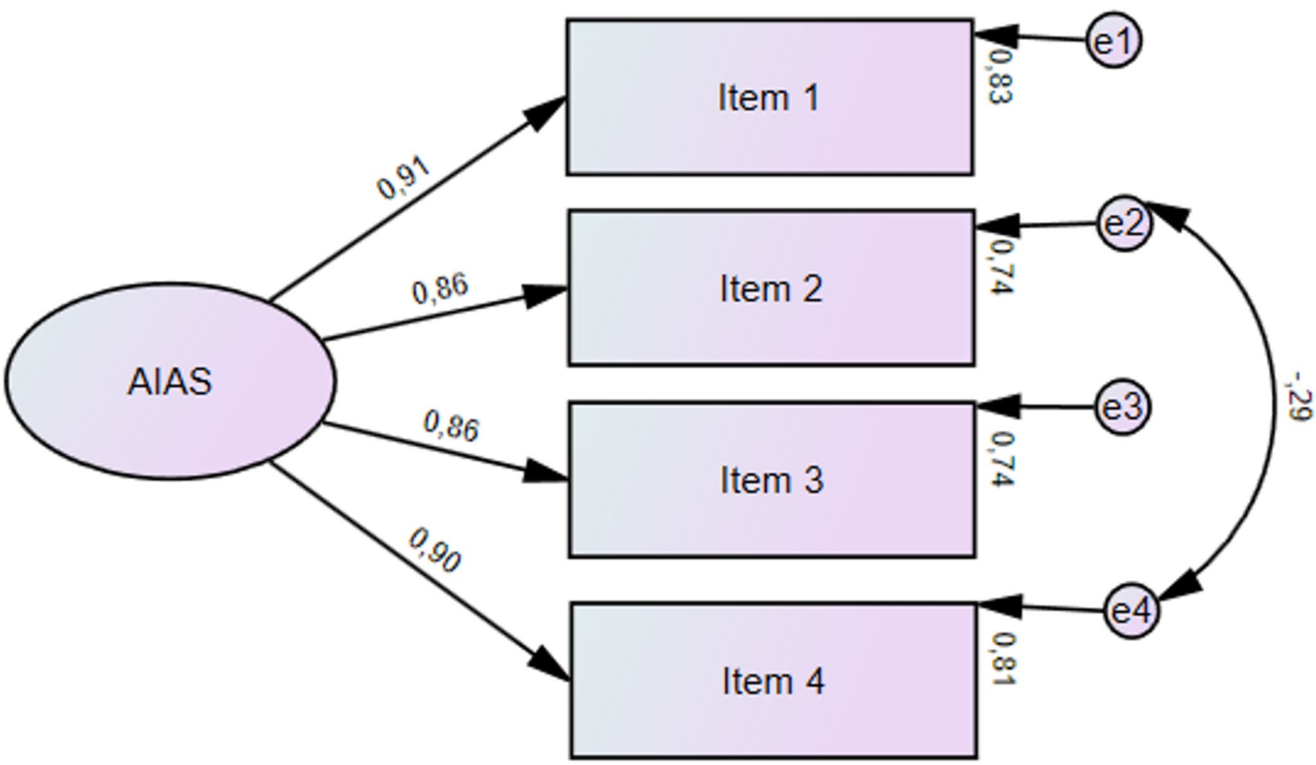

Fig. 1 Structural factors of the AIAS-4 in Polish sample

(N=511), British (N=230), and American (N=300) populations. Data from the British and American samples was the same used in the scale validation study (Grassini, 2023). Please consult the validation study method section for details about the data collection.

Firstly, descriptive statistics were calculated for each scale item in every country (refer to Table 3).

After confirming the baseline model, which consisted of one factor with four items as described above, we conducted a multi-group confirmatory factor analysis (MGCFA) to investigate the factorial invariance and latent mean differences. Bootstrapped ML was used due to the non-normality of the distribution, which was confirmed by multivariate kurtosis statistics (c.r=5.036).

To assess the equivalence of measurement, we included additional indicators to evaluate the model's fit to the data (Byrne, 2016). We considered a decrease in CFI of less than or equal to −0.010 and a change in RMSEA of less than or equal to 0.015, as well as changes in AIC parameters. A change of more than 1 allows us to infer equivalence, following Zercher et al. (2015). The calculations were performed using SPSS 28 and AMOS.

First, a baseline model with no equality constraints across groups was tested to assess configurational invariance (MCI). At this stage, the model is referred to as the configurational model and is tested against the results of all participants from Poland, Great Britain, and United States of America. The configural invariant model showed very good fit to the data, as indicated by CFI=1.00; RMSEA=0.000 (90% CI [0.000, 0.030]; AIC=78.855 (Table 5). Such an exceptionally good model fit may reflect not only the parsimony of the model itself but also the influence of the estimation procedure (ML with bootstrapping), which—through repeated resampling—may produce stabilized standard error estimates and slightly upward-biased fit statistics.

Next, the Metric Invariance Model (MMI) was tested, in which all factor loadings were constrained to be equal across groups. The metric equivalence indices were satisfactory except for the RMSEA indicator, which exceeded the recommended cut-off point>0.015. Hence, partial metric equivalence was estimated in the subsequent stage. The fit of the model to the data was assessed sequentially by removing the imposed restrictions of equality of factor loadings on successive items. It was found that releasing item 4 resulted in partial metric equivalence: The decrease in all fit rate of the partial metric model compared to the configural model does not exceed the assumed cut-off points (Table 4).

Finally, the last model– Scalar Invariance Model (MSI) was tested in which item intercepts were constrained to be equal across groups. The decrease in fit rate of the scalar model compared to the partial metric model does exceed all the assumed cut-off points: the **Δ**CFI was above the 0.010 cut-off (0.025 as well as ΔRMSEA was above 0.015 **[**0,06]). Therefore, it can be concluded that scalar equivalence has not been confirmed (Table 4).

Table 3 Descriptive statistics and psychometric properties for AIAS items in Poland (N=511), the UK (N=230) and the USA (N=300)

		Min	Max	M	SD	Skewness	Kurtosis	Discriminatory power	Cronbach's alpha
PL	S_1	1	10	5.35	2.404	−0.216	−0.557	0.88	
	S_2	1	10	4.98	2.551	−0.008	−0.803	0.80	
	S_3	1	10	5.83	2.493	−0.365	−0.618	0.83	0.931
	S_4	1	10	5.66	2.414	−0.299	−0.519	0.84	
UK	S_1	1	10	5.06	2.275	−0.114	−0.622	0.83	
	S_2	1	10	5.09	2.424	−0.061	−0.686	0.80	
	S_3	1	10	6.19	2.330	−0.322	−0.364	0.78	0.902
	S_4	1	10	5.57	1.990	0.099	−0.017	0.73	
USA	S_1	1	10	5.80	2.252	−0.131	−0.435	0.87	
	S_2	1	10	5.53	2.537	−0.057	−0.758	0.78	
	S_3	1	10	7.14	2.361	−0.552	−0.368	0.73	0.906
	S_4	1	10	6.17	2.200	−0.271	−0.234	0.79	

Table 4 Assessment of measurement equivalence levels between Polish (N=511), British (N=230), and American (N=300) populations

	df	χ2	χ2/df	RMSEA	LO	Hi	CFI	AIC	Δχ2 (df)
MCI	3	0.855	0.285	0.000	0.000	0.030	1.000	78.855	
MMI	9	17.694	1.966*	0.031	0.007	0.051	0.997	83.694	16.839 (6)**
PMMI (item 4 released	7	7.351	1.050	0.007	0.000	0.039	1.000	77.351	6.496(4)
MSI	17	94.526	5.560***	0.067	0.054	0.080	0.975	144.526	71.363 (10)***

MCI model of configural invariance, *MMI* model of metric invariance, *PMMI* partial model of metric invariance, *MSI* model of scalar invariance

*** p<.001; ** p<.01; *** p<.05

Discussion

The aim of this study was to assess the measurement equivalence of the AIAS-4 in three distinct countries: Poland, the UK, and the USA. Prior to assessing the equivalence of the measurement, the structure factor in Poland was evaluated. Both statistical analyses– EFA and CFA– confirmed that the 1-factor solution is the best fit to the data, thus confirming that the baseline model is identical in Poland, the UK and the USA.

Multi-group confirmatory factor analysis (MGCFA) confirms the configural measurement of the AIAS-4, indicating that the models have the same structure and comparable patterns of association between the factor loadings of the statements and the latent factors of the scale in all groups (configural invariance). The construct of general attitudes towards artificial intelligence is understood similarly by respondents from Poland, the UK, and the US. The difference relates to one claim of the scale (item 4), as indicated by the results of the partial metric equivalence of the scale. It means that the test results of individuals from different countries with the same level of the trait being tested (general positive attitude to AI) will be the same, except item 4. This item is about the importance of AI in general for humanity (*I think AI technology is positive for humanity*), in contrast to the other three statements which focus on the personal impact of artificial intelligence on improving one's own life (*I believe that AI will improve my life*), work (*I believe that AI will improve my work*), and personal decision to use AI in the future (*I think I will use AI technology in the future).*

The three statements pertain to the respondents' intentions or advantages of using AI in their work and personal lives. In contrast, statement 4 addresses social and cultural factors, as well as varying interpretations of what is positive for humanity in the three countries where the study was conducted. Previous research has shown that cultural differences have an effect on how people perceive such type of social positive attitudes toward AI (Grassini & Ree, 2023).

The partial metric equivalence, with the exclusion of item 4, indicates that although respondents demonstrate configural equivalence in their understanding of attitudes towards AI, they nevertheless judge it differently in terms of AI's importance to humanity (item 4). The consequence of this is that item 4 should be excluded from the scale in order to facilitate cross-cultural research. An alternative approach would be to modify this claim, although this should be preceded by emic-type studies in which indicators, for example, scale claims measuring a given construct, are developed separately in each culture analysed (Lubiewska & Glogowska, 2018). In the context of the research presented here, it would be necessary to conduct country-specific analyses to ascertain the understanding of the term 'the meaning of AI for humanity' among representatives of different countries.

There is no confirmation of scalar equivalence among cross-cultural study groups. This means that comparisons between these countries in terms of differences relating to general attitudes towards AI should be avoided. Partial metric equivalence allows us to estimate the relationship between attitudes towards AI and other variables, but item 4 should be excluded (Cheung & Rensvold, 2002; Schmitt & Kuljanin, 2008; Van De Schoot et al., 2012).

The confirmation of scalar equivalence in cross-cultural studies is rendered challenging when the samples under consideration exhibit discrepancies in their responses to the questionnaires administered. This is exemplified in the research presented here, particularly with regard to item 4, whose equivalence was not confirmed even at the stage of metric equivalence. This indicates that the respondents from the compared countries, despite having a similar understanding of the measured construct and attitudes towards AI, assess these attitudes differently, which is undoubtedly influenced by culture and the associated degree of technology maturity and development.

To our knowledge, this is the first study to investigate the measurement invariance of the AIAS-4 across different countries. The confirmation of configural equivalence indicates that, in all comparisons, attitudes towards AI are understood in a similar manner; however, the measurement of this variable differs with respect to Item 4. Given that the other three items refer to the personal use of AI and Item 4 to its relevance to humanity, it may be posited that we are dealing with two independent attitudes. In future research, it may be of value to separate these sub-constructs, namely the personal and humanity-wide relevance of AI.

These findings contribute to our understanding of cross-cultural differences in attitudes towards AI by highlighting how cultural contexts influence perceptions of AI's role both personally and on a societal level. The partial metric invariance observed, particularly in responses to AI's impact on humanity (item 4), suggests that societal and cultural factors shape how individuals view AI's broader implications.

Limitations

The main limitation of the presented research is the cross-sectional nature of the study– we cannot state the scale's accuracy when assessing constructs over time. As a consequence, future research should investigate AIAS-4 longitudinal invariance.

Another limitation is the sample diversity within the countries studied. While the study includes respondents from Poland, the UK, and the USA, it may not fully capture the wide range of cultural, socioeconomic, and demographic backgrounds within these countries, that all represent western type societies with a developed economy and high standard of living. Future research could address this by including a broader and more diverse sample to examine.

The practical implications of the research presented include the possibility of assessing differences of averages in attitudes towards AI across countries, but separately. Confirmation of metric equality - with the exception of item 4 - allows for correlational analyses and comparisons between countries. A practical conclusion is also the prospect of future research on a method to measure attitudes towards AI, which - as suggested by the results obtained - should perhaps separate attitudes towards AI into two dimensions: personal meaning (for the individual) and general human meaning (for humanity).

Appendix

AIAS-4 (PL) (Polska adaptacja: W. Talik, E. Talik, S. Grassini)

Oceń, w jakim stopniu zgadzasz lub nie zgadzasz się z poniższymi twierdzeniami. Swoją oceną wyraź na skali od 1 do 10, gdzie 1 oznacza, że w ogóle NIE zgadzasz się z tym stwierdzeniem, a 10, że zgadzasz się z nim całkowicie.

Uważam, że sztuczna inteligencja poprawi moje życie.	W ogóle się nie zgadzam	1	2	3	4	5	6	7	8	9	10	Całkowicie się zgadzam
Uważam, że sztuczna inteligencja usprawni moją pracę.	W ogóle się nie zgadzam	1	2	3	4	5	6	7	8	9	10	Całkowicie się zgadzam
Myślę, że będę korzystać z technologii sztucznej inteligencji w przyszłości.	W ogóle się nie zgadzam	1	2	3	4	5	6	7	8	9	10	Całkowicie się zgadzam
Myślę, że technologia sztucznej inteligencji jest pozytywna dla ludzkości/ przynosi ludzkości korzyści.	W ogóle się nie zgadzam	1	2	3	4	5	6	7	8	9	10	Całkowicie się zgadzam

Author contributions Wiesław Talik– Conceptualization, Methodology, Investigation, Resources, Data Curation, Writing - Original Draft, Writing - Review & Editing, Visualization, Project administration, Funding acquisition.
Elżbieta Talik– Methodology, Formal analysis, Resources, Writing - Original Draft, Writing- Review & Editing.
Simone Grassini– Conceptualization, Investigation, Resources, Data Curation, Writing- Original Draft, Writing - Review & Editing.

Funding This research was partially supported by Grant nr 1118 from Institute of Social Science, The John Paul II Catholic University of Lublin.

Data availability The authors have no financial or proprietary interests in any material discussed in this article.

Declarations

All authors certify that they have no affiliations with or involvement in any organization or entity with any financial interest or non-financial interest in the subject matter or materials discussed in this manuscript.

Publication ethics All procedures were in accordance with the Declarationof Helsinki. Informed consent was obtained from all individual adult participants included in the study.

Competing interests The authors have no competing interests to declare that are relevant to the content of this article.

Conflict of interest The authors have no relevant financial or non-financial interests to disclose.

References

Brislin, R. W. (1976). Comparative research methodology: Cross-cultural studies. *International Journal of Psychology, 11*(3), 215–229. https://doi.org/10.1080/00207597608247359

Byrne, B. M. (2016). Testing instrument equivalence across cultural groups: Basic concepts, testing strategies, and common complexities. In N. Zane, G. Bernal, & F. T. L. Leong (Eds.), *Evidence-based psychological practice with ethnic minorities: Culturally informed research and clinical strategies* (pp. 125–143). American Psychological Association. https://doi.org/10.1037/14940-007

Byrne, B. M., Shavelson, R. J., & Muthén, B. (1989). Testing for the equivalence of factor covariance and mean structures: The issue of partial measurement invariance. *Psychological Bulletin*, *105*(3), 456–466. https://doi.org/10.1037/0033-2909.105.3.456

Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling: A Multidisciplinary Journal, 9*(2), 233–255. https://doi.org/10.1207/S15328007SEM0902_5

Cieciuch, J., & Davidov, E. (2015). Establishing measurement invariance across online and offline samples. A tutorial with the software packages Amos and Mplus. *Studia Psychologica: Theoria Et Praxis*, *15*(2), 83–99. https://doi.org/10.21697/sp.2015.14.2.06

Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly*, *13*(3), 319–340.

Douglas, S. P., & Craig, C. S. (2007). Collaborative and iterative translation: An alternative approach to back translation. *Journal of International Marketing*, *15*(1), 30–43. https://doi.org/10.1509/jimk.15.1.030

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

Grassini, S., & Ree, A. S. (2023). Hope or doom AI-ttitude? Examining the impact of gender, age, and cultural differences on the envisioned future impact of artificial intelligence on humankind. In *Proceedings of the European Conference on Cognitive Ergonomics 2023* (pp. 1–7).

Hofstede, G. (1984). *Culture's consequences: International differences in work-related values* (Vol. 5). Sage.

Hofstede, G. (2011). Dimensionalizing cultures: The Hofstede model in context. *Online Readings in Psychology and Culture*, *2*(1), 8.

Hornowska, E., & Paluchowski, W. J. (2004). Kulturowa adaptacja testów psychologicznych. [W:] J. Brzeziński (red.), Metodologia badań naukowych– wybór tekstów (s. 151–191). Warszawa: PWN.

Huang, C., Zhang, Z., Mao, B., & Yao, X. (2022). An overview of artificial intelligence ethics. *IEEE Transactions on Artificial Intelligence*. https://doi.org/10.1109/TAI.2022.3194503

Jan, J., Alshare, K. A., & Lane, P. L. (2024). Hofstede's cultural dimensions in technology acceptance models: A meta-analysis. *Universal Access in the Information Society*, *23*(2), 717–741. https://doi.org/10.1007/s10209-022-00930-7

Kostick-Quenet, K. M., Cohen, I. G., Gerke, S., Lo, B., Antaki, J., Movahedi, F., Njah, H., Schoen, L., Estep, J. E., & Blumenthal-Barby, J. S. (2022). Mitigating racial bias in machine learning. *Journal of Law, Medicine & Ethics*, *50*(1), 92–100. https://doi.org/10.1017/jme.2022.13

Kummer, T. F., & Schmiedel, T. (2016). Reviewing the role of culture in strategic information systems research: A call for prescriptive theorizing on culture management. *Communications of the Association for Information Systems*, *38*(1), 5. https://doi.org/10.17705/1CAIS.03805

Lubiewska, K., & Głogowska, K. (2018). Zastosowanie analizy Równoważności Pomiarowej w Badaniach Psychologicznych. *Polskie Forum Psychologiczne*, *23*(2), 330–356. https://doi.org/10.14656/PFP20180207

Neudert, L. M., Knuutila, A., & Howard, P. N. (2020). *Global Attitudes Towards AI, Machine Learning & Automated Decision Making.* Working paper 2020.10 Oxford, UK: Oxford Commission on AI & Good Governance, 10pp.

Rodrigues, R. (2020). Legal and human rights issues of AI: Gaps, challenges and vulnerabilities. *Journal of Responsible Technology*, *4*, 100005. https://doi.org/10.1016/j.jrt.2020.100005

Schepman, A., & Rodway, P. (2023). The general attitudes towards artificial intelligence scale (GAAIS): Confirmatory validation and associations with personality, corporate distrust, and general trust. *International Journal of Human-Computer Interaction*, *39*(13), 2724–2741. https://doi.org/10.1080/10447318.2022.2085400

Schmitt, N., & Kuljanin, G. (2008). Measurement invariance: Review of practice and implications. *Human Resource Management Review*, *18*(4), 210–222. https://doi.org/10.1016/j.hrmr.2008.03.003

Van De Schoot, R., Lugtig, P., & Hox, J. (2012). A checklist for testing measurement invariance. *European Journal of Developmental Psychology*, *9*(4), 486–492. https://doi.org/10.1080/17405629.2012.686740

van de Vijver, F., & Hambleton, R. K. (1996). Translating tests: Some practical guidelines. *European Psychologist*, *1*(2), 89–99. https://doi.org/10.1027/1016-9040.1.2.89

Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. D. (2003). User acceptance of information technology: Toward a unified view. *MIS Quarterly*, *27*(3), 425–478.

Vincent-Lancrin, S., Van der Vlies R. (2020). Trustworthy artificial intelligence (AI) in education: Promises and challenges. *OECD Education Working Papers* 218 1–17.

Webster, C., & Ivanov, S. (2019). Robotics, artificial intelligence, and the evolving nature of work. In *Springer eBooks* (pp. 127–143). https://doi.org/10.1007/978-3-030-08277-2_8

Yu, C. Y., & Muthén, B. (2002). Evaluating cutoff criteria of model fit indices for latent variable models with binary and continuous outcomes. Doctoral Dissertation. http://www.statmodel.com/download/Yudissertation.pdf

Zercher, F., Schmidt, P., Cieciuch, J., & Davidov, E. (2015). The comparability of the universalism value over time and across countries in the European social survey: Exact vs. approximate measurement invariance. *Frontiers in Psychology*, *6*. https://doi.org/10.3389/fpsyg.2015.00733

Open Science The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.