Current Psychology (2025) 44:12315–12327
https://doi.org/10.1007/s12144-025-08009-1

Artificial Intelligence Attitudes Inventory (AIAI): development and validation using Rasch methodology

Christian U. Krägeloh[1] · Vladimir Melekhov[2] · Mohsen M. Alyami[3] · Oleg N. Medvedev[4]

Accepted: 24 May 2025 / Published online: 10 June 2025

Abstract

With the rapid advancements in artificial intelligence (AI), it is vital to develop psychometrically sound measures of public attitudes toward this technology. The present study aimed to refine a pool of candidate items to create a concise yet robust inventory for assessing attitudes toward AI. A total of 96 items—drawn from reworded robot-related scales and newly developed items to reflect AI-specific themes—were administered to a sample of 604 adults from the general population of the United States (age range: 18–89 years; 48% male). Iterative Rasch analysis was used to reduce the number of items while ensuring psychometric robustness, applying multiple criteria for item selection including fit residuals, differential item functioning (DIF), and conceptual clarity. The resulting scale, named the Artificial Intelligence Attitudes Inventory (AIAI), consists of two 8-item subscales measuring positive and negative attitudes toward AI. Analyses revealed that these subscales are distinct constructs rather than opposites on a single continuum, and they are only weakly related to psychological distress. The AIAI provides a concise yet comprehensive measure of positive and negative attitudes toward AI that can be efficiently administered alongside other measures. The findings underscore the multifaceted nature of public perceptions of AI and highlight the need for further research into the profiles and determinants of these attitudes. As AI continues to shape our world, the AIAI offers a valuable tool for understanding and monitoring public sentiment toward this transformative technology.

Keywords Artificial intelligence · Attitudes · AIAI, psychometrics · Rasch analysis

In recent years, the rapid advancement of artificial intelligence (AI) has marked a significant milestone in the history of technology, capturing the attention of both the public and the scientific community. The emergence of AI language models like ChatGPT has sparked widespread interest and discourse, as people grapple with the potential impact of this technology on various aspects of human life (Taecharungroj, 2023). Unlike previous technological advancements that primarily influenced physical tasks and manual labour (Parsons, 1985), AI tools like ChatGPT are poised to transform "white-collar" work by executing complex tasks such as problem-solving, language translation, writing, and data analysis (Haque, 2022; Taecharungroj, 2023).

Attitudes toward such transformative technologies are not easily reducible to simple approval or disapproval. Theoretically, attitudes are understood as latent predispositions to evaluate an object in a favorable or unfavorable manner, inferred from individuals' responses such as verbal judgments, behavioral tendencies, or physiological reactions (Schwarz, 2000). Importantly, attitudinal responses are often context-dependent and shaped by the interplay of cognitive, affective, and social factors. Dual-process models further suggest that positive and negative attitudes may function as partially independent dimensions, rather than opposite ends of a single continuum (Cacioppo & Berntson, 1994). This conceptualization is especially relevant in the context of AI, where excitement about technological progress frequently coexists with ethical concerns and fears

✉ Christian U. Krägeloh
chris.krageloh@aut.ac.nz

[1] Department of Psychology and Neuroscience, School of Clinical Sciences, Auckland University of Technology, 90 Akoranga Drive, Northcote, Auckland 1142, New Zealand

[2] University of Auckland, Auckland, New Zealand

[3] Department of Psychology, College of Education, University of Ha'il, Ha'il, Saudi Arabia

[4] School of Psychological and Social Sciences, University of Waikato, Hamilton, New Zealand

of disruption. Therefore, instruments designed to measure public attitudes toward AI must be theoretically grounded and sensitive to such multidimensionality.

Despite these conceptual nuances, there remains a gap in the availability of psychometrically robust tools designed to capture the complexity of public sentiment toward AI. The swift advancement of AI, coupled with its heightened visibility in the media, has only recently necessitated the development of methods to gauge public attitudes toward AI. This is in stark contrast to the field of robotics, for which a wide range of acceptability and attitude questionnaires have been available for quite some time (Krägeloh et al., 2019). One of these questionnaires is the 14-item Negative Attitudes towards Robots Scale (NARS; Nomura et al., 2006), which Persson et al. (2021) re-worded for AI contexts and named Negative Attitudes towards Artificial Intelligence Scale (NAAIS). An exploratory factor analysis found that the NAAIS is best interpreted as a three-factor profile expressing negative attitudes toward (1) concrete use of AI, (2) hypothetical use of AI, and (3) emotions and relationships with AI.

Another questionnaire to measure AI attitudes was developed by Schepman and Rodway (2020). Schepman and Rodway (2022) reported on the results of a confirmatory factor analysis of this tool, the General Attitudes towards Artificial Intelligence Scale (GAAIS). This 20-item instrument measures positive and negative attitudes toward AI expressed as separate factors. More recently, Stein et al. (2024) introduced the Attitudes Toward AI and Robotics Index (ATTARI-12), a 12-item questionnaire that was also designed to measure attitudes toward AI, independent of specific contexts. Unlike the GAAIS, the ATTARI-12 is a unidimensional scale, incorporating cognitive, affective, and behavioral facets into a single measure to provide a comprehensive yet succinct overview of an individual's attitude toward AI.

In addition to these more comprehensive questionnaires, there are also some notably shorter instruments designed to quickly capture specific dimensions of AI attitudes. The 5-item Attitude Towards Artificial Intelligence (ATAI) scale scores attitudes as a two-factor profile (Sindermann et al., 2021). This scale includes two items assessing acceptance—specifically, the extent to which participants trust AI and believe it will benefit humanity. The remaining three items measure fear, addressing concerns about AI's potential to cause job losses and even the destruction of humankind. Furthermore, the AI Attitudes Scale (AIAS-4; Grassini, 2023) offers another succinct measure, interpreted as a single factor. This scale includes items that gauge whether respondents believe AI will improve their personal lives and work environments. It also assesses the perceived likelihood of using AI technology in the future and whether this technology is viewed as positive for humanity.

The availability of various instruments to measure attitudes toward AI is encouraging and reflects the growing interest and complexity of this field. Very brief scales like the 5-item ATAI (Sindermann et al., 2021) and the 4-item AIAS-4 (Grassini, 2023) are particularly useful for quick assessments but lack the breadth needed to tap into more detailed themes. Additionally, very brief questionnaires provide limited range in differentiation across participants, thus making them vulnerable to floor and ceiling effects. The 14-item NAAIS (Persson et al., 2021), while more robust in this regard, drew exclusively from items of one single robot attitudes scale and thus may not capture themes specific to AI (e.g., "The word 'robot' means nothing to me" or "I would feel very nervous just standing in front of a robot"). The 20-item GAAIS (Schepman & Rodway, 2022), in contrast, captures a broad perspective of both positive and negative attitudes but contains some items that are very nonspecific (e.g., "Artificial Intelligence is exciting") or items that may be expressing anxiety more than attitudinal aspects (e.g., "I shiver with discomfort when I think about future uses of Artificial Intelligence"). Lastly, the ATTARI-12 (Stein et al., 2024) integrates affective, behavioral, and cognitive aspects into a single measure. While this may help understand how these aspects are interrelated in forming attitudes, these distinctions are not reflected in the factor structure, and these aspects are integrated into a single score.

Despite the variety of available instruments, no existing measure fully meets the dual requirement of psychometric rigor and conceptual clarity specific to AI. Very brief tools such as the ATAI (Sindermann et al., 2021) and AIAS-4 (Grassini, 2023) are limited in scope and prone to floor or ceiling effects. The NAAIS (Persson et al., 2021), though more comprehensive, relies exclusively on reworded items from a robot-focused measure and may not adequately capture uniquely AI-related themes. While the GAAIS (Schepman & Rodway, 2022) offers a broad coverage of attitudes, it includes items that may reflect anxiety or arousal rather than evaluative judgment. Similarly, the ATTARI-12 (Stein et al., 2024) integrates multiple facets of attitudes into a single factor score, limiting interpretability across dimensions. Thus, there remains a need for a validated instrument that distinctly measures both positive and negative attitudes toward AI, grounded in current theoretical perspectives and developed through contemporary psychometric techniques.

The purpose of the present study was to develop a concise yet comprehensive measure of general AI attitudes by validating candidate items specifically tailored for assessing attitudes toward AI, distinct from anxiety-related responses. Krägeloh et al. (2024) recently created a set of candidate items for a new AI attitudes questionnaire by drawing in a wide range of robot-related themes but also themes specific to AI. The authors utilized modern AI tools like ChatGPT to re-word existing items from two robot-related questionnaires, namely the above-mentioned 14-item NARS and the 30-item and Frankenstein Syndrome Questionnaire (FSQ; Nomura et al., 2012). The FSQ was designed to assess attitudes toward humanoid robots, specifically addressing anxieties and ethical concerns around their integration into society. This questionnaire captures themes such as trust, interpersonal fears, and principles regarding humanoid robots. In addition to drawing on the NARS and FSQ, Krägeloh et al. (2024) used a multiple-step process (including the use of AI for item generation, checking for redundancy, and rephrasing for clarity and relevance) to generate additional candidate itemsthat reflect concerns and perceptions specific about AI. This systematic approach ensured that 44 items drawn from robot-related questionnaires were supplemented by further 52 items to address specific contemporary issues associated with AI technology, such as ethical considerations and potential societal impacts (Krägeloh et al., 2024). Utilization of this large pool of candidate items maximized the likelihood that a sufficiently large of set of items was available from which to select the psychometrically robust items that also capture relevant aspects of AI attitudes.

To ensure psychometric rigor, the present study utilized Rasch analysis to examine the scale's internal structure, evaluate item functioning, and assess measurement precision. Compared to traditional psychometric approaches, Rasch models enable the development of shorter yet more reliable scales by testing item fit, threshold ordering, and invariance across population groups (Medvedev & Krägeloh, 2022; Tennant & Conaghan, 2007). Rasch analysis also to test for unidimensionality, addresses local response dependencies, and supports person–item targeting on a common latent trait continuum. These features were deemed critical for the construction of a theoretically sound, psychometrically robust measure of attitudes toward AI. In addition, the present study explored the relationship between the new measure and key demographic variables, as well as its correlation with non-clinical measures of depression, anxiety, and stress. This was to ensure that the new tool provides a measure of attitudes only, as it is important that such measures are not confounded by affect (Brown et al., 2015). By establishing a psychometrically sound and efficient measure of AI attitudes and examining its correlates, this study aimed to contribute to a better understanding of public sentiment toward AI and the ongoing dialogue on the integration of AI into society. Measuring public attitudes towards technology is crucial for fostering a transparent and balanced dialogue, as attitudes often guide behavior and influence the adoption and perception of new technologies in society (Kerschner & Ehlers, 2016).

Method

Participants

All participants were recruited using Qualtrics to be a representative sample of adults living in the United States. While the total sample consisted of 609 participants, the responses from 5 participants were excluded from the analysis as their completion time for the survey exceeded 2 h. The age range was from 18 years to 89 years, with M= 46.75 and SD= 18.21. Of the 604 participants, 292 (48.3%) identified as male, 311 (51.5%) as females, and 1 (0.2%) as "Other". In terms of ethnicity, 370 participants (61.3%) identified as White, 145 (24.0%) identified as Black, and 14.7% selected "Other". Slightly more than half of the sample (n= 331, 54.8%) indicated that their highest education level was high school education or less, 168 (27.8%) had a university undergraduate degree, and 105 (17.4%) had a postgraduate degree. Participants were located in the "South" of the United States (n= 240, 39.7%), 134 (22.2%) from the "Midwest", 111 (18.4%) from the "Northeast", and 119 (19.7%) from the "West".

Procedure

Participants were recruited online using convenience sampling via Qualtrics. Prior to completing the questionnaire, participants provided electronic informed consent. No incentives were provided to the participants. Data were collected in January 2024, and the study had been approved by the authors' university ethics review board.

Measures

The 96 candidate items for an AI acceptability or attitudes scale proposed by Krägeloh et al. (2024) were used as potential items for a final scale to be developed. This included (1) 30 items drawn from the FSQ (Nomura et al., 2012), such as "I am afraid that humanoid robots will make us forget what it is like to be human" re-worded as "I am afraid that AI will make us forget what it is like to be human", (2) 14 items drawn from the NARS (Nomura et al., 2006), such as "I would feel uneasy if I was given a job where I had to use robots", re-worded as "I would feel uneasy if I was given a job where I had to use AI", and (3) newly worded items such as "I believe that AI can contribute positively to the creative industries, such as art, music, and writing" and "I am worried that AI could lead to a loss of privacy for individuals". Items were thus a combination of positive and negative aspects about AI. Although no formal cognitive interviews or structured expert panel review were conducted, item development was informed by established robot attitudes measures and AI-specific themes drawn from the contemporary literature, with conceptual clarity considered during item refinement alongside psychometric criteria. Responses were provided on a 5-point Likert scale, with the following response options: 1 = *Not agree at all*, 2 = *Slightly agree*, 3 = *Moderately agree*, 4 = *Agree to a larger extent*, to 5 = *Completely agree*.

The brief version of the Depression Anxiety Stress Scales (DASS-21; Lovibond & Lovibond, 1995) includes three 7-item subscales assessing depression, anxiety, and stress in non-clinical populations. Participants were instructed to read each statement and indicate to what extent it applied to them over the past week. The response options for each item ranged from 0 to 3 (0 = *Did not apply to me at all* to 3 = *Applied to me very much*, *or most of the time*). Total subscale scores were calculated by summing the scores of the relevant items. Following the scale developers' recommendation, all scores were multiplied by 2 to make scores comparable to those of the full 42-item version (Lovibond & Lovibond, 1995).

Data analyses

All descriptive analyses and preliminary psychometric analyses were conducted using SPSS v. 28. Due to the set-up of the online questionnaire, all questions needed to be completed, which meant that there were no missing data. In preparation for Rasch analysis, age was re-categorized into the following three approximately equal sized categories: 18–35 years (n = 202), 36–57 years (n = 204), and 58 to 89 years (n = 198). Because of severe under-representation in the gender category "Other", only two gender categories ("male" and "female") were retained.

Because the Rasch model assumes a unidimensional scale, the negative attitude items were reverse scored so that their psychometric properties could be examined alongside the positively worded items. The negatively worded items were initially identified based on face validity but also aided by a principal components analysis (PCA) that forced a single component. Once 54 items had been identified as representing negative attitudes and had been recoded, an unconstrained PCA with oblique rotation (promax) was conducted. This analysis resulted in an uninterpretable solution with a highly complex cross-loading pattern across 11 components. Forcing a three-component solution still revealed an uninterpretable solution. Forcing a solution with two components aligned items according to their direction of wording (positive versus negative), albeit with substantial cross-loadings that suggested that a one-component solution is most tenable. This result confirmed that this set of 96 items with 54 re-coded negatively worded items was adequate to be analyzed using Rasch analysis. The re-scoring of negatively worded items is indicated here through the addition "r", such as Item 58r.

Rasch analysis was conducted using the software package RUMM2030 (Andrich et al., 2009). The sample size exceeded 500 and thus met the requirements to achieve robust item calibration and stable person measures (Linacre, 1994). Prior to the primary analysis, a likelihood-ratio test was conducted, which confirmed the appropriateness of using the unrestricted partial credit model of the Rasch approach ($p < 0.001$) as suggested by Masters (1982). The model should show non-significant overall and individual Chi-square fit statistics, adjusted by Bonferroni for significance at $p < 0.05$. Item fit residuals are considered elevated if they fall outside the range of −2.50 to + 2.50. Items with fit residuals exceeding |2.5| were thus considered misfitting, in line with standard guidelines for Rasch modeling (Tennant & Conaghan, 2007), and were iteratively deleted to improve model fit. The residual correlation matrix must not show signs of local dependency among items; a correlation higher than 0.20 relative to the average residual correlation suggests such dependency (Christensen et al., 2016). Combining dependent items into subtests can address this issue (Lundgren-Nilsson et al., 2013; Wainer & Kiely, 1987). For example, when subgroups of items showed common stems or phrasing (e.g., "I believe…"), subtests were created to address potential method effects.Additionally, significant differential item functioning (DIF) should not be present due to personal factors (in this case, gender, age group, educational category, and region of residence). The person separation index (PSI), akin to Cronbach's alpha, reflects the precision with which subjects are distributed along the measurement continuum and checks the reliability of subscales in Rasch analysis (Tennant & Conaghan, 2007).

In assessing dimensionality in Rasch analysis, independent-samples *t*-tests are utilized to compare person estimates between two item groups. This method focuses on the highest positive and negative factor loadings from the first principal component of residuals after removing the latent factor (Smith, 2002). If less than 5% of these *t*-tests are significant, or if the 5% threshold overlaps with the lower boundary of a binomial confidence interval calculated for the number of significant *t*-tests, the scale is regarded as unidimensional (Tennant & Pallant, 2006).

An initial Rasch analysis incorporating all 96 items necessitated the iterative removal of 82 items. Throughout each iteration, items exhibiting the highest degree of misfit were systematically excluded. This process continued until a subset of 14 items demonstrated an acceptable fit. Notably, this subset consisted exclusively of positively worded items. This pattern suggested the presence of a method effect (as also noted in through the PCA with a forced two-component solution), where negatively worded items were progressively identified as misfits and subsequently eliminated. As a result, each iterative step increasingly predisposed the remaining negatively worded items to misfit, leading to their removal. Consequently, the final item set was unbalanced, containing only positively worded items, which does not provide a comprehensive representation of attitudes. Due to these limitations, this approach was ultimately deemed unsuitable. Therefore, an alternative strategy was adopted, involving separate Rasch analyses for the 42 positively worded items and the 54 negatively worded items, to ensure a more balanced evaluation of the scale.

Results

Even though PSI was excellent (PSI = 0.96), the initial baseline model of the 42 positive items exhibited a significant item-trait interaction ($\chi^2(378)$ = 2243.92, $p < 0.01$) and was thus an inadequate fit to the Rasch model. The relevant fit statistics are shown in Table 1, and the detailed item location information is presented in Table S1 found in the Supplementary Materials. The items with the clearly most extreme fit residuals (Item 15: 12.99, Item 38: 14.63, Item 40: 14.43, Item 41: 11.21, Item 85: 13.48, and Item 90: 11.97) were deleted, and the model was re-run. Significant item misfits exceeding the threshold of |2.50| continued to occur in subsequent iterations. Initially, items with fit residual exceeding |5.00| were discarded, which resulted in the deletion of Items 28, 31, 33, 53, 71, and 74, followed by deletion of Items 11, 26, 56, and 79 in a subsequent iteration. The fit of the model with the remaining 26 items was still significant ($\chi^2(234)$ = 352.14, $p < 0.01$), prompting the need for further item deletion. Over the course of further iterations, items with a fit residual exceeding |3.00| were deleted. This resulted in the exclusion of Items 2, 6, 13, 19, 20, 21, 22, 23, 24, 43, and 49. The resulting fit with 15 items continued to have excellent internal consistency (PSI = 0.93), but the fit was still significant ($\chi^2(135)$ = 166.08, $p < 0.01$). In the iterations that followed, further items were deleted as a result of elevated item fit residuals (Items 46, 54, 64, 66, and 68), which reduced the number of positively worded candidate items to 10.

Even though this 10-item solution continued to demonstrate excellent internal consistency (PSI = 0.91), the item-trait interaction was still significant ($\chi^2(90)$ = 124.58, $p <$ 0.01). At this stage, there were no more misfitting items, and additional criteria were considered for item deletion. Based on significant DIF by age, Item 68 was discarded. Specifically, using a Bonferroni-adjusted significance level ($p < 0.01$), DIF analysis revealed that respondents in the oldest age group (58–89 years) answered this item differently compared to those in the youngest group (18–35 years). After accounting for the same overall level of positive attitudes, the youngest group generally endorsed Item 68 more strongly than the oldest group.

Considering the remaining nine items, it was evident that Item 82 ("AI can create new forms of interactions both between humans and between humans and machines") stood out as not being indicative of either positive and negative attitudes toward AI and instead appeared to express a

Table 1 The overall Rasch model fit statistics (item fit residuals, person fit residuals, goodness of fit) for the separate fits of the positive and negative attitude items. Also shown are person separation index (PSI) and the results of the unidimensionality test

	Item fit residual		Person fit residual		Goodness of fit		PSI	Significant *t*-tests	
Analyses	Value/SD		Value/SD		χ2 (df)	*p*		%	Lower bound %
<u>Positive items</u>									
Pos1 (42 items)	1.28	5.72	−0.67	3.60	2243.92 (378)	< 0.01	0.96	11.09	9.35
Pos2 (26 items)	0.94	2.09	−0.78	2.94	352.14 (234)	< 0.01	0.95	8.61	6.87
Pos3 (15 items)	0.67	1.51	−0.82	2.62	166.08 (135)	< 0.01	0.93	3.48	1.74†
Pos4 (10 items)	0.59	1.62	−0.72	2.11	124.58 (90)	< 0.01	0.91	3.15	1.41†
Pos5 (8 items)	0.66	1.65	−0.68	1.92	112.18 (72)	< 0.01	0.89	2.81	1.08†
Pos6 (8 items with subtest)	0.33	2.94	−0.62	1.52	51.35 (45)	0.24	0.87	0.99	−0.74†
<u>Negative items</u>									
Neg1 (54 items)	1.34	4.86	−0.71	3.86	2152.17 (486)	< 0.01	0.97	15.73	13.99
Neg2 (51 items)	1.25	4.12	−0.69	3.66	1763.71 (459)	< 0.01	0.97	16.06	14.32
Neg3 (34 items)	1.25	2.54	−0.73	3.24	662.31 (306)	< 0.01	0.95	8.11	6.37
Neg4 (21 items)	0.99	1.79	−0.64	2.57	355.88 (189)	< 0.01	0.93	6.46	4.72†
Neg5 (10 items)	0.76	1.74	−0.62	2.05	130.18 (90)	< 0.01	0.88	1.99	0.25†
Neg6 (8 items)	0.49	1.49	−0.59	1.84	110.81 (72)	< 0.01	0.86	0.50	−1.24†
Neg7 (8 items with subtest)	−1.07	7.00	−0.61	1.29	30.87 (27)	0.28	0.85	1.16	−0.58†

†Unidimensionality confirmed based on results from Smith's test (2000)

fact statement. Based on this conceptual consideration, this item was deleted. The fit of the remaining 8-item solution was still significant ($\chi^2(72)$ = 112.18, $p < 0.01$). In a final iteration of the analysis, a subtest was created to control for a potential method effect due to common item wording. Items 80, 81, 86, and 94 all started with the phrase "I believe…". Combining these items into a subtest resulted in a non-significant fit ($\chi^2(45)$ = 51.35, $p > 0.05$). PSI was 0.87, thus exceeding the threshold of 0.85 for the suitability of the instrument for within-subject comparisons (Tennant & Conaghan, 2007). There was no significant DIF, and Smith's (2002) found no evidence to suggest that this final 8-item subscale was other than unidimensional (with a percentage of significant *t*-test of 1% and thus below the 5-% criterion). The person-item distribution (top panel of Fig. 1) indicates that, when omitting extreme responses, the majority of the participants' levels of the latent trait (i.e., positive attitudes toward AI) are covered by the items (as indicated in the blue bars). There are slight floor and ceiling effects, although with less than 5% of respondents at each end of the continuum.

The analysis of the 54 negatively worded candidate items followed the same analysis approach. The fit statistics are shown in Table 1, and the item locations for the

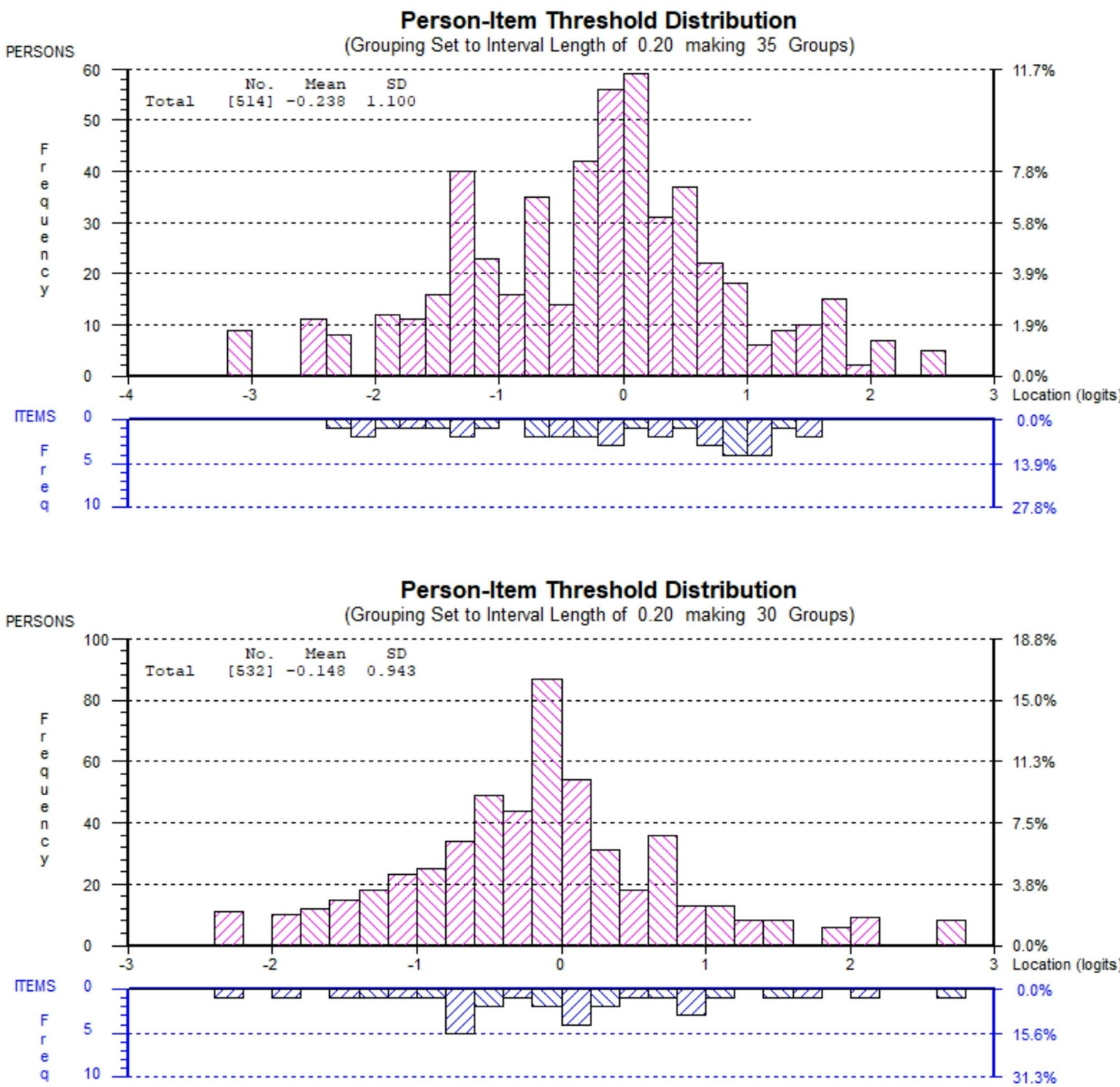

Fig. 1 Person-item distribution for the 8-item subscale assessing positive attitudes toward AI (top panel) and negative attitudes (bottom panel)

baseline model for the negative items are displayed in Table S2 (Supplementary Materials). The initial baseline model (Neg1) yielded a significant item-trait interaction ($\chi^2(486) = 2152.17, p < 0.01$) with PSI = 0.97. Deletion of three highly misfitting items (Item 32r: 13.40, Item 76r: 12.14, and Item 77r: 16.92) resulted in a model fit that was still significant ($\chi^2(459) = 1763.71, p < 0.01$). Subsequent iterations deleted items with fit residual exceeding |5.00|, which were Items 1r, 3r, 5r, 7r, 12r, 16r, 25r, 35r, 36r, 42r, 50r, 55r, 60r, 61r, 87r, 88r, and 89r. The resulting 34-item model continued to be significant ($\chi^2(306) = 662.31, p < 0.01$). Further iterations from deletion of items with fit residuals above |3.00| (Items 8r, 9r, 10r, 14r, 17r, 37r, 52r, 62r, 63r, 65r, 67r, 69r, and 93r) resulted in a 21-item model that had a PSI of 0.93 but a significant item-trait interaction ($\chi^2(189) = 355.88, p < 0.01$). In the subsequent iterations, items with elevated fit residuals (>|2.50) continued to be an issue, prompting the deletion of Items 4r, 18r, 27r, 30r, 59r, 70r, 72r, 84r, 91r, 92r, and 96r. The resulting 10-item solution had a PSI of 0.88. Even though item-trait interaction was still significant ($\chi^2(90) = 130.18, p < 0.01$), there were no remaining elevated item fit residuals that could prompt the deletion of further items.

As with the positively worded items, conceptual criteria were then considered to inform any further analysis steps. Consideration of Item 39r ("I feel that if we become over-dependent on AI, something bad might happen") and Item 51r ("I feel that if I depend on AI too much, something bad might happen") needed to be deleted. This is because the wording of "over-dependent" and "too much" and linking those to "something bad might happen" result in a trivial statement that would be difficult not to agree with. Deletion of these two items for a resulting 8-item version still resulted in a significant fit ($\chi^2(72) = 110.81, p < 0.01$). However, controlling for a method effect by combining all items starting with the stem "I am…" (Items 29r, 34r, 44r, 47r, 48r, and 58r), the fit was no longer significant ($\chi^2(27) = 30.87, p > 0.05$). PSI was 0.85 and thus met the criterion for reliable within-participant analyses. There was no significant DIF by person factors, and the model can be considered unidimensional. The person-item distribution (bottom panel of Fig. 1), with extreme responses omitted, shows that all participants' levels of the latent trait (i.e., negative attitudes toward AI) are covered by the items (as indicated in the blue bars).

As the Rasch analysis indicated that the 8-item subscale representing positive attitudes toward AI and the 8-item subscale about negative attitudes can be used with strong validity and reliability, subscale summary scores were generated. The positive subscale and negative subscale (with negative items still re-scored) were weakly correlated ($r = 0.29, p < 0.01$), indicating that these were distinct construct rather than simple opposites. A correlation matrix indicated that items within the positive subscale were moderately to highly correlated, with Spearman's ρ-coefficients ranging from 0.54 to 0.73. Similarly, coefficients for correlation of items within the negative subscale ranged from 0.45 to 0.68. Cronbach's α and McDonald's ω for the positive subscale were both 0.93, with all items correlated with the item total by at least 0.73. No item deletion increased Cronbach's α and McDonald's ω. For the negative subscale, Cronbach's α and McDonald's ω were both 0.91, again with no increase in internal consistency if any item was deleted. Item-to-total correlation coefficients for the negative items were no less than 0.67. A different picture emerged when correlating positive items with negative items. Here, Spearman's ρ-coefficients ranged from 0.09 ($p < 0.05$, Items 94 and 58r) to 0.26 ($p < 0.01$, Items 78 and 73r). These results indicate that the two subscales are indeed distinct and that the scores should not be interpreted as summary profile.

Because the negatively worded items were no longer compared with positively worded items, the reverse scoring was no longer applied. Table 2 shows an overview of the final 8-item positive subscale with their new item labels (P1 to P8) and the final labels for the negative subscale (N1 to N8), together with item means and SD. There were no significant gender differences for the positive ($t(601) = 1.51, p > 0.05$) and negative ($t(601) = -0.36, p > 0.05$) subscales. An analysis of variance (ANOVA) revealed no effect by ethnicity for the positive subscale ($F(2, 601) = 2.55, p > 0.05$) but an effect for the negative subscale ($F(2, 601) = 8.24, p < 0.01$). A Tukey post-hoc analysis indicated that, compared to Black participants ($M = 23.42$, $SD = 7.64$), White participants ($M = 26.78$, $SD = 9.08$) had significantly stronger negative views of AI ($p < 0.01$). Furthermore, an ANOVA showed a significant effect of education level for the positive subscale ($F(2, 601) = 16.62, p < 0.01$). All group comparisons were significant. Participants with high school education or less had the lowest mean ($M = 20.49$, $SD = 8.60$), followed by participants with a university undergraduate degree ($M = 22.76$, $SD = 9.02$) and those with a postgraduate degree ($M = 26.12$, $SD = 9.63$). In terms of negative attitudes, there was no effect by education level ($F(2, 601) = 0.30, p > 0.05$). For region of residence, there was no effect on positive attitude levels ($F(2, 601) = 1.56, p > 0.05$) or negative attitudes ($F(2, 601) = 1.42, p > 0.05$).

Table 3 shows Pearson's r coefficients (all $p < 0.01$) for correlations between the positive and negative subscales with age and the three DASS-21 subscales. As re-scoring was no longer applied, the two subscales were now negatively correlated ($r = 0.29, p < 0.01$). All correlations were weak. Age was negatively correlated with positive attitudes and positively with negative attitudes. Both subscales were positively correlated with depression, anxiety, and stress.

Table 2 Overview of the final Artificial Intelligence Attitudes Inventory (AIAI) items by subscale, with item means and *SD*

Item number	Item wording	Mean	SD
Positive attitudes subscale			
P1	I trust AI to make fair and transparent decisions that affect me or others.	−0.27	0.49
P2	I like the idea that AI can augment and enhance human intelligence and creativity.	−0.18	0.59
P3	I am interested in learning more about AI and its applications for my personal and professional development.	−0.35	1.16
P4	I believe that AI can help reduce social inequalities by providing better access to education and healthcare.	−0.30	−1.53
P5	I believe that AI can help empower individuals by providing them with personalized services and support.	−0.04	12.99
P6	Persons and organizations related to the development of AI are well-meaning.	0.28	−0.09
P7	I believe that AI can help us achieve a more sustainable and eco-friendly future.	0.22	−2.00
P8	I believe that AI has the potential to improve the quality and accessibility of healthcare services.	0.17	−1.09
Negative attitudes subscale			
N1	I am concerned that AI will harm humanity and society.	0.22	−1.31
N2	I am afraid that AI will make us forget what it is like to be human.	−0.08	−5.05
N3	I am afraid that AI will encourage less interaction between humans.	−0.15	−3.19
N4	Something bad might happen if AI developed into living beings.	0.40	2.12
N5	I am worried that AI could lead to a loss of privacy for individuals.	0.37	6.46
N6	I am concerned that AI may be used for malicious purposes, such as cyber-attacks or surveillance.	0.12	4.05
N7	I am concerned that AI might be used to manipulate public opinion or spread misinformation.	0.15	6.27
N8	I would hate the idea that AI was making judgments about things.	−0.10	14.63

Note. that the original wording for P8 was “I believe AI has the potential to improve the quality and accessibility of healthcare services”. For consistency with question stems for items in this subscale, the question stem for Item P8 was changed to “I believe that…”

Table 3 Pearson’s *r* correlation coefficients for associations between positive AI attitudes, negative attitudes, age, and the three DASS-21 subscales depression, anxiety, and stress. All coefficients were significant at $p < 0.01$

	Positive attitudes to AI	Negative attitudes to AI	Age	Depression	Anxiety
Negative attitudes to AI	−0.29	-			
Age	−0.20	0.20	-		
Depression	0.20	0.17	−0.32	-	
Anxiety	0.29	0.12	−0.37	0.84	-
Stress	0.24	0.20	−0.35	0.86	0.86

Discussion

The purpose of the present study was to provide a psychometric analysis of previously developed candidate items (Krägeloh et al., 2024) for a scale designed to measure attitudes towards AI. While this list of 96 potential items included a comprehensive coverage of relevant themes drawn from robot attitudes and acceptability scales, namely the NARS and FSQ, more than half of the items were new creations to tap into a wide range of additional AI-specific themes. Given the size of this initial item pool, a key aim was to refine and streamline it, reducing the burden on respondents and ensuring its psychometric robustness. To achieve this, we employed a modern Rasch analysis approach, analyzing responses from a large and diverse U.S. sample. Iterative Rasch analysis resulted in a scale that we introduce here as the Artificial Intelligence Attitudes Inventory (AIAI). The AIAI includes a positive attitudes subscale and a negative attitudes subscale, with 8 items each.

Our initial Rasch analysis attempted to create a unidimensional scale by re-scoring negatively worded items and exploring to what extent they expressed the same concept as the positively worded ones. Both Rasch analysis and classical test theory approaches (including PCA) revealed that sources of model misfit cannot be attributed to method effects from item wording only and that the negative attitudes items were not simply expressing the opposite of positive items. The AIAI thus does not apply re-scoring of the negatively worded items since the items are not combined into a single scale but are instead scored as separate measures of positive and negative attitudes toward AI.

The fact that positive and negative attitudes toward AI were only weakly negatively correlated may indicate that these aspects – at least as expressed by the content of the AIAI items – appear to be somewhat decoupled and thus only in a weak reciprocal relationship. According to the evaluative activation modes framework discussed by Cacioppo and Berntson (1994), positive and negative evaluative processes can operate independently or even simultaneously under certain conditions. In the context of AI, it is possible that individuals may recognize and appreciate the benefits and advancements brought by AI technology (positive attitudes) while simultaneously harboring concerns about issues such as privacy, autonomy, or job displacement (negative attitudes). This decoupling may be indicative of the complex and multifaceted nature of public perceptions toward AI, where traditional bipolar models of attitude measurement might not fully capture the nuanced evaluative dimensions, people consider when forming their views. Moreover, the notion that positive and negative attitudes cannot simply be viewed as opposites on a spectrum is further supported by research in other domains. For example, Pittinsky et al. (2011) have demonstrated in their studies on attitudes toward minority groups that positive and negative attitudes can exist as theoretically and functionally distinct constructs. This underscores the complexity of attitudinal processes and suggests that similar distinctions may be relevant when considering attitudes toward technological entities like AI.

While it is clear that positive and negative attitudes toward AI, as measured by the AIAI, are not completely decoupled—given the presence of a weak negative correlation—this relationship invites further investigation into the various profiles of attitudes that may exist among different individuals. Future research could profitably employ methods such as latent profile analysis (Bravo et al., 2016) to identify distinct groups of respondents for whom these attitudes may be decoupled or, conversely, more tightly interwoven. The direction of the relationship of positive and negative attitudes toward AI and robots may certainly also directly depend on the type of questionnaire used and how items are worded. For instance, a study of participants in India using the FSQ to assess robot acceptability found that positive and negative attitudes were weakly positively correlated (Bharatharaj et al., 2022), unlike the negative correlation found with the AIAI. Future investigations will need to examine whether variations in the type of AI scale used can lead to different outcomes in terms of the coupling or decoupling of these attitudes.

The fact that the three subscales of the DASS-21, namely Depression, Anxiety, and Stress, were only weakly correlated with the AIAI subscales indicates that the AIAI can suitably be described as a scale to measure AI attitudes as opposed to AI anxiety. Not much information is currently available about the extent to which other existing AI attitude questionnaires are correlated with psychological distress. In the development of the ATAI (Sindermann et al., 2021), ATTARI-12 (Stein et al., 2024), GAAIS (Schepman & Rodway, 2022), and NAAIS (Persson et al., 2021), no such measures were included. The only except was the AIAS-4 (Grassini, 2023). Here, the scale developer correlated the 4-item AIAS ("I believe that AI will improve my life", "I believe that AI will improve my work", "I think I will use AI technology in the future", and the reverse-scored item "I think AI technology is a threat to humans") with the Media and Technology Usage and Attitudes Scale (MTUAS; Rosen et al., 2013) and found that the AIAS was weakly correlated ($r = 0.25$) with the MTUAS subscale measuring anxiety related to the absence or unavailability of technology such as phone or internet.

Of note in the results of the present study is the fact that the DASS-21 subscales were weakly positively correlated with both positive as well as negative attitudes subscales of the AIAI. With age, in contrast, there was a weakly negative correlation with positive attitudes and a weakly positive correlation with negative attitudes. Again, this testifies to the complex interplay of the factors contributing to attitudes about AI. For age, therefore, attitudes can be described as being in reciprocal negative activation (Cacioppo & Berntson, 1994). For individuals experiencing heightened psychological distress, on contrast, these attitudes could be described to be co-activating. Individuals with psychological distress might be overly vigilant about AI developments, or individuals with low distress levels may not be overly concerned about AI and thus have no strong feelings about AI, whether it be positive or negative. This pattern aligns with theoretical models suggesting that affective arousal – often associated with psychological distress – can intensify cognitive and affective processing across both positive and negative evaluative dimensions, rather than simply biasing judgements toward negativity (Storbeck & Clore, 2008). Several theoretical mechanisms may explain this co-activation pattern. Individuals with psychological distress might exhibit heightened vigilance and attentional broadening (Cisler & Koster, 2010) to potential impacts of technologies like AI, making them more attuned to both benefits and risks simultaneously. Similarly, the uncertainty management model (van den Bos, 2009) suggests that individuals experiencing distress may be more motivated to seek information and form opinions about novel phenomena with uncertain outcomes—such as AI—leading to amplified engagement across both positive and negative dimensions. Alternatively, this pattern could reflect a mood-congruent processing effect as described by Forgas (2008), where mild negative affect promotes more careful, systematic processing of complex information, resulting in recognition of both positive and negative aspects rather than simplified evaluations. Individuals with low distress levels, by contrast, may not be overly concerned about AI and thus have no strong feelings about AI, whether positive or negative. This theoretical framing helps explain why psychological distress may correlate positively with both attitude dimensions rather than simply shifting attitudes in a single direction.

In the development of the candidate items for the AIAI, 44 items were drawn from the robot-related FSQ and NARS, and 52 items were items specifically created to tap into AI-related issues. In final 8-item positive attitudes subscale of the AIAI, only one item (P6) of the FSQ was retained, which means that the remaining items were specifically created for AI contexts. It appears, therefore, that the opportunities and benefits that participants attribute to AI are very different to those attributed to robots. It is probably not unreasonable to claim that this is expected because most individuals would have had more chances to have memorable encounters with AI as opposed to with a robot.

For the 8-item negative attitudes subscale of the AIAI, in contrast, three items were drawn from the FSQ (N2, N3, and N8) and one item from the NARS (N8). This distinction in item retention between the subscales might reflect how popular media, like in *I*, *Robot* (Asimov, 2004), emphasize the dangerous characteristics of robots through the fact that they are linked to AI, especially fears related to robotic autonomy influenced by AI decision-making. Future research should further explore this hypothesis, possibly using mixed-method approaches to understand these distinctions better. As robotic technology evolves and becomes more integrated into daily life, attitudes toward robots and AI might become more similar. The intertwined nature of robots and AI is expressed in the definition of the field of robopsychology (Krägeloh et al., 2022). As this connection between robots and AI is increasingly recognized, future attitude scales might therefore not be limited to one or the other but instead explore attitudes to robots and AI simultaneously.

With the rapid advancements in AI technology, one can certainly expect increasing interest in the measurement of AI attitudes. While several tools have already been developed, the AIAI can be described as having distinct advantages. This includes the fact that the AIAI drew on a wide range of item content, including those specifically developed for AI context and not only items adapted from existing robot attitudes or anxiety scales (e.g., NAAIS; Persson et al., 2021). Unlike some of the very short measures such as the 4-item AIAS-4 (Grassini, 2023) and the 5-item ATAI (Sindermann et al., 2021), the AIAI assesses a larger range of positive and negative attitudes with two subscales of 8 items each. While the AIAI is structurally similar to the GAIS (Schepman & Rodway, 2022), it contains some items that are nonspecific and broad (e.g., "Artificial Intelligence is exciting"), offering limited informative value regarding the facets that contribute to positive or negative attitudes. Moreover, some item content in the GAIS may lean more toward measuring anxiety rather than purely attitudinal aspects (e.g., "I shiver with discomfort when I think about future uses of Artificial Intelligence"), which risks confounding attitude measurement with affective responses (Brown et al., 2015). The development of the AIAI minimized the use of language that refers to strong emotional reactions. This way, researchers can explore the associations between AI attitudes and psychological variables such as psychological distress without the risk of spurious correlations arising from content overlap in the measurement instruments.

When comparing the psychometric properties of the AIAI with other published AI attitude scales, several strengths emerge. The AIAI's rigorous development through Rasch analysis represents a methodological advantage over other scales, which typically relied on classical test theory approaches alone. This modern psychometric approach enabled thorough detailed examination of item functioning, which are analyses not reported for other AI attitude measures. Furthermore, the AIAI explicitly demonstrated that positive and negative attitudes represent distinct constructs rather than opposite ends of a continuum, a psychometric insight not explored in most other measures. The AIAI is also unique in its verified distinction from anxiety, with weak correlations to psychological distress variables providing discriminant validity evidence.

The present study has several psychometric limitations that should be considered when interpreting results. While the AIAI demonstrated excellent internal consistency (Cronbach's α and McDonald's ω of 0.93 for the positive subscale and 0.91 for the negative subscale), temporal stability was not assessed. The absence of test-retest reliability data limits our understanding of whether AIAI scores represent stable traits or are susceptible to situational influences. Our validation approach primarily established structural validity through Rasch analysis. However, several important validity aspects remain unaddressed. We did not examine relationships with behavioral criteria related to AI acceptance or resistance (criterion validity), comparisons with other established AI attitude measures were not conducted (convergent validity), and relationships with conceptually distinct constructs beyond psychological distress variables were not examined (discriminant validity). The present study explored the relationship with psychological distress to verify that items are indeed about attitudes and not about anxiety. However, due to response burden from completing the long 96-item pool of candidate items and the 21-item DASS-21, no other measures were included.

While the AIAI provides a robust framework for measuring attitudes towards AI, there are also several other limitations to the study that must be considered. Firstly, the data used to validate the AIAI came exclusively from a single cultural context, the United States. This limitation is significant as attitudes toward technology, including AI, can vary widely across different cultural settings. Future studies should look to assess the cross-cultural generalizability of the AIAI, exploring whether certain items or themes carry different weights or meanings in different countries. Such exploration could benefit from employing DIF analysis across diverse cultural contexts. Secondly, the sample was a convenience sample recruited through Qualtrics, which may limit the generalizability of the findings even within the United States. Participants from convenience panels may not fully represent the broader population in terms of demographic diversity, technological familiarity, or attitudinal patterns.

Future research should seek to replicate the findings using more representative sampling methods. Additionally, expanding the scale's use to other languages, similar to how the ATAI (Sindermann et al., 2021) is available in multiple languages, would enhance its applicability and reach. Future research could also explore the relationship of the AIAI with personality variables such as reported by Stein et al. (2024). Specifically, future studies could examine the association between AIAI scores and specific personality traits such as conscientiousness and neuroticism, given their established links with technology usage (Barnett et al., 2015). Some work has started to explore the effect of personality variables on attitudes toward AI (Park & Woo, 2022). Future work could explore the interplay of a range of personal factors in relation to positive and negative attitudes toward AI using approaches such as network analysis (Chalmers et al., 2022). The advantage of the AIAI is that it provides a comprehensive measure of AI attitudes with a fairly small number of items to permit concomitant collection of additional measures.

Cross-cultural validation should not only involve linguistic translations but also include comparisons between populations that differ in their levels of AI integration, for instance, contrasting countries with widespread AI use against those in emerging markets. Additionally, it would be insightful to explore how the AIAI is related to the other existing measures of AI attitudes and also additional background information about the participants such as personal experience with AI or extent of media consumption that is about AI. Lastly, in order to explore the stability of the construct, future work should conduct test-retest reliability or use the more comprehensive approach of Generalizability Theory, which is able to differentiate clearly between dynamic and stable aspects in self-report scales (Medvedev et al., 2017).

To conclude, the present study successfully refined and validated the AIAI using a comprehensive psychometric approach, including Rasch analysis. From an initial pool of 96 items—some drawn from existing scales like the FSQ and NARS and newly created items to address AI-specific themes—the AIAI emerged with distinct subscales for positive and negative attitudes. As with the growing number of scales available to measure robot acceptability and attitudes (David et al., 2022; Krägeloh et al., 2019), increases in number of scales for AI attitudes is an advantage as it permits researchers to select the most appropriate tool for their purposes. Some of this development has resulted in instruments for very specific purposes such as attitudes toward AI at work (Park et al., 2024) and in military contexts (Hadlington et al., 2023). The AIAI, in contrast, was developed to provide a measure of a range of general positive and negative attitudes to be used in diverse contexts. The present study demonstrated that this measure has very strong psychometric properties. The AIAI thus provides a practical and robust tool for researchers and policymakers interested in tracking public attitudes toward AI. Its two subscales allow for the independent assessment of positive and negative sentiments, which can inform the design of public education campaigns, communication strategies, or interventions aimed at promoting informed and balanced perceptions of AI technologies.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s12144-025-08009-1.

Funding Open Access funding enabled and organized by CAUL and its Member Institutions

References

Andrich, D., Sheridan, B., & Luo, G. (2009). *RUMM 2030*. RUMM Laboratory.

Asimov, I. (2004). *I, robot* (Vol. 1). Spectra.

Barnett, T., Pearson, A. W., Pearson, R., & Kellermanns, F. W. (2015). Five-factor model personality traits as predictors of perceived and actual usage of technology. *European Journal of Information Systems*, *24*(4), 374–390. https://doi.org/10.1057/ejis.2014.10

Bharatharaj, J., Kutty, S. K. S., Munisamy, A., & Krägeloh, C. U. (2022). What do members of Parliament in India think of robots? Validation of the Frankenstein syndrome questionnaire and comparison with other population groups. *International Journal of Social Robotics*, *14*(9), 2009–2018. https://doi.org/10.1007/s12369-022-00921-x

Bravo, A. J., Boothe, L. G., & Pearson, M. R. (2016). Getting personal with mindfulness: A latent profile analysis of mindfulness and psychological outcomes. *Mindfulness*, *7*(2), 420–432. https://doi.org/10.1007/s12671-015-0459-7

Brown, G. P., Hawkes, N., Cooper, A., Jonsdottir, S., & Tata, P. (2015). The anxiety attitude and belief Scale-2: Development, measurement model, and initial validity. *Clinical Psychology & Psychotherapy*, *22*(6), 687–697. https://doi.org/10.1002/cpp.1928

Cacioppo, J. T., & Berntson, G. G. (1994). Relationship between attitudes and evaluative space: A critical review, with emphasis on the separability of positive and negative substrates. *Psychological Bulletin*, *115*(3), 401–423. https://doi.org/10.1037/0033-2909.115.3.401

Chalmers, R. A., Cervin, M., & Medvedev, O. N. (2022). Network analysis. In O. N. Medvedev, C. U. Krägeloh, R. J. Siegert, & N. N. Singh (Eds.), *Handbook of assessment in mindfulness research*. Springer. https://doi.org/10.1007/978-3-030-77644-2_70-1

Christensen, K. B., Makransky, G., & Horton, M. (2016). Critical values for yen's Q3: Identification of local dependence in the Rasch model using residual correlations. *Applied Psychological Measurement*, *41*(3), 178–194. https://doi.org/10.1177/0146621616677520

Cisler, J. M., & Koster, E. H. (2010). Mechanisms of attentional biases towards threat in anxiety disorders: An integrative review. *Clinical Psychology Review*, *30*(2), 203–216. https://doi.org/10.1016/j.cpr.2009.11.003

David, D., Thérouanne, P., & Milhabet, I. (2022). The acceptability of social robots: A scoping review of the recent literature. *Computers in Human Behavior*, *137*, 107419. https://doi.org/10.1016/j.chb.2022.107419

Forgas, J. P. (2008). Affect and cognition. *Perspectives on Psychological Science*, *3*(2), 94–101. https://doi.org/10.1111/j.1745-6916.2008.00067.x

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.2023.1164810

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

Haque, M. A. (2022). A brief analysis of ChatGPT– a revolutionary tool designed by openai. *EAI Endorsed Transactions on AI and Robotics*, *1*(1), e15. https://doi.org/10.4108/airo.v1i1.2983

Kerschner, C., & Ehlers, M. H. (2016). A framework of attitudes towards technology in theory and practice. *Ecological Economics*, *126*, 139–151. https://doi.org/10.1016/j.ecolecon.2016.02.010

Krägeloh, C. U., Bharatharaj, J., Kutty, S. K. S., Nirmala, P. R., & Huang, L. (2019). Questionnaires to measure acceptability of social robots: A critical review. *Robotics*, *8*(4), 88. https://doi.org/10.3390/robotics8040088

Krägeloh, C. U., Bharatharaj, J., Albo-Canals, J., Hannon, D., & Heerink, M. (2022). The time is ripe for robopsychology. *Frontiers in Psychology*, *13*, 968382. https://doi.org/10.3389/fpsyg.2022.968382

Krägeloh, C. U., Alyami, M. M., & Medvedev, O. N. (2024). AI in questionnaire creation: Guidelines illustrated in AI acceptability instrument development. In C. U. Krägeloh, M. Alyami, & O. N. Medvedev (Eds.), *International handbook of behavioral health assessment*. Springer. https://doi.org/10.1007/978-3-030-89738-3_62-1

Linacre, J. M. (1994). Sample size and item calibration [or person measure stability. *Rasch Measurement Transactions*, *7*(4), 328.

Lovibond, S. H., & Lovibond, P. F. (1995). *Manual for the depression anxiety stress scales* (2nd ed.). Psychology Foundation of Australia.

Lundgren-Nilsson, Å., Jonsdottir, I. H., Ahlborg, G., & Tennant, A. (2013). Construct validity of the psychological general well being index (PGWBI) in A sample of patients undergoing treatment for stress-related exhaustion: A Rasch analysis. *Health and Quality of Life Outcomes*, *11*, 2. https://doi.org/10.1186/1477-7525-11-2

Masters, G. N. (1982). A Rasch model for partial credit scoring. *Psychometrika*, *47*(2), 149–174. https://doi.org/10.1007/BF02296272

Medvedev, O. N., & Krägeloh, C. U. (2022). Rasch measurement model. In O. N. Medvedev, C. U. Krägeloh, R. J. Siegert, & N. N. Singh (Eds.), *Handbook of assessment in mindfulness research*. Springer. https://doi.org/10.1007/978-3-030-77644-2_4-1

Medvedev, O. N., Krägeloh, C. U., Narayanan, A., & Siegert, R. J. (2017). Measuring mindfulness: Applying generalizability theory to distinguish between state and trait. *Mindfulness*, *8*(4), 1036–1046. https://doi.org/10.1007/s12671-017-0679-0

Nomura, T., Suzuki, T., Kanda, T., & Kato, K. (2006). Measurement of negative attitudes toward robots. *Interaction Studies*, *7*, 437–454. https://doi.org/10.1075/is.7.3.14nom

Nomura, T., Sugimoto, K., Syrdal, S. S., Dautenhahn, K. (2012). Social acceptance of humanoid robots in Japan: A survey for development of the Frankenstein Syndorome [sic] Questionnaire. In Proceedings of the 12th IEEE-RAS International Conference on Humanoid Robots, Osaka, Japan, 29 November–1 December 2012; pp. 242–247. https://doi.org/10.1109/HUMANOIDS.2012.6651527

Park, J., & Woo, S. E. (2022). Who likes artificial intelligence? Personality predictors of attitudes toward artificial intelligence. *Journal of Psychology*, *156*(1), 68–94. https://doi.org/10.1080/00223980.2021.2012109

Park, J., Woo, S. E., & Kim, J. (2024). Attitudes towards artificial intelligence at work: Scale development and validation. *Journal of Occupational and Organizational Psychology*. https://doi.org/10.1111/joop.12502

Parsons, H. M. (1985). Automation and the individual: Comprehensive and comparative views. *Human Factors*, *27*(1), 99–111. https://doi.org/10.1177/001872088502700109

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

Pittinsky, T. L., Rosenthal, S. A., & Montoya, R. M. (2011). Liking is not the opposite of disliking: The functional separability of positive and negative attitudes toward minority groups. *Cultural Diversity and Ethnic Minority Psychology*, *17*(2), 134–143. https://doi.org/10.1037/a0023806

Rosen, L. D., Whaling, K., Carrier, L. M., Cheever, N. A., & Rokkum, J. (2013). The media and technology usage and attitudes scale: An empirical investigation. *Computers in Human Behavior*, *29*, 2501–2511. https://doi.org/10.1016/j.chb.2013.06.006

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-Computers Interaction*, *39*(13), 2724–2741. https://doi.org/10.1080/10447318.2022.2085400

Schwarz, N. (2000). Attitude measurement. In A. E. Kazdin (Ed.), *Encyclopedia of psychology* (Vol. 1, pp. 313–316). American Psychological Association. https://doi.org/10.1037/10516-109

Sindermann, C., Sha, P., Zhou, M., Wernicke, J., Schmitt, H. S., Li, M., Sariyska, R., Stavrou, M., Becker, B., & Montag, C. (2021). Assessing the attitude towards artificial intelligence: Introduction of a short measure in german, chinese, and english Language. *KI-Künstliche Intelligenz,35*(1), 109–118. https://doi.org/10.1007/s13218-020-00689-0

Smith, E. V. (2002). Detecting and evaluation the impact of multidimensionality using item fit statistics and principal component analysis of residuals. *Journal of Applied Measurement*, *3*(2), 205–231.

Stein, J. P., Messingschlager, T., Gnambs, T., Hutmacher, F., & Appel, M. (2024). Attitudes towards AI: Measurement and associations with personality. *Scientific Reports*, *14*, 2909. https://doi.org/10.1038/s41598-024-53335-2

Storbeck, J., & Clore, G. L. (2008). Affective arousal as information: How affective arousal influences judgments, learning, and memory. *Social and Personality Psychology Compass*, *2*(5), 1824–1843. https://doi.org/10.1111/j.1751-9004.2008.00138.x

Taecharungroj, V. (2023). What can ChatGPT do? Analyzing early reactions to the innovative AI chatbot on Twitter. *Big Data and Cognitive Computing*, *7*(1), 35. https://doi.org/10.3390/bdcc7010035

Tennant, A., & Conaghan, P. G. (2007). The Rasch measurement model in rheumatology: What is it and why use it? When should it be applied, and What should one look for in a Rasch paper? *Arthritis & Rheumatism*, *57*(8), 1358–1362. https://doi.org/10.1002/art.23108y

Tennant, A., & Pallant, J. F. (2006). Unidimensionality matters! (a Tale of two Smiths?). *Rasch Measurement Transactions*, *20*, 1048–1051.

van den Bos, K. (2009). Making sense of life: The existential self trying to deal with personal uncertainty. *Psychological Inquiry*, *20*(4), 197–217. https://doi.org/10.1080/10478400903333411

Wainer, H., & Kiely, G. L. (1987). Item clusters and computerized adaptive testing: A case for testlets. *Journal of Educational Measurement*, *24*(3), 185–201.

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