# Measuring Attitudes Towards Artificial Intelligence (ATTARI): Introducing the ATTARI-6 and ATTARI-3 Short Forms

Markus Appel[1], Timo Gnambs[2], Tanja V. Messingschlager[1], Fabian Hutmacher[1], & Jan-Philipp Stein[3]

[1] University of Würzburg, Würzburg, Germany

[2] Leibniz Institute for Educational Trajectories, Bamberg, Germany

[3] Chemnitz University of Technology, Chemnitz, Germany

## Author Note

Markus Appel http://orcid.org/0000-0003-4111-1308

Timo Gnambs https://orcid.org/0000-0002-6984-1276

Tanja V. Messingschlager https://orcid.org/0000-0001-6737-8997

Fabian Hutmacher https://orcid.org/0000-0002-0018-2559

Jan-Philipp Stein https://orcid.org/0000-0003-3874-0277

**Accepted to be published in the journal *Telematics and Informatics***

We have no conflicts of interest to disclose. Studies 1a and 1b were not preregistered. Study 2 was preregistered at https://aspredicted.org/tx23c2.pdf. The data and analysis code are provided at https://osf.io/byq6n.

Correspondence concerning this article should be addressed to Markus Appel, University of Würzburg, Oswald-Külpe-Weg 82, 97074 Würzburg. E-mail: markus.appel@uni-wuerzburg.de.

**Abstract**

The results of three samples are presented to introduce two short forms of the Attitudes Towards Artificial Intelligence Scale (ATTARI). These short forms consist of six items (ATTARI-6) and three items (ATTARI-3), respectively, that reflect cognitive, emotional, and behavioral attitude facets. The participants ($N_{total}$ = 3,171) worked on the English or German language scales. The results corroborate high reliability of the short forms as well as strong factorial validity and measurement invariance. Correlations and non-inferiority tests suggest that the short forms yield associations with related variables that are similar in size to correlations observed for the full scale. Several fields of application for the short forms are discussed.

*Keywords*: artificial intelligence; attitudes; short scale; survey

# Measuring Attitudes Towards Artificial Intelligence (ATTARI): Introducing the ATTARI-6 and ATTARI-3 Short Forms

Artificial intelligence (AI) has rapidly emerged as one of the most transformative forces of the 21st century. From using large language models and related platforms such as ChatGPT as a source for news and information to medical diagnostics and algorithmic decision-making, AI is increasingly embedded into everyday life (e.g., Maslej et al., 2025). Its diffusion has sparked both enthusiasm and apprehension: Whereas many highlight AI's potential to increase efficiency, foster innovation, and solve complex societal problems, others warn of risks such as job displacement, algorithmic bias, surveillance, and the erosion of human creativity (Malecki et al., 2025; Vorobeva et al., 2022; Weinmann et al., 2026). As AI systems have become pervasive, public acceptance and trust in these technologies will be pivotal to their responsible adoption, making the study of attitudes toward AI a pressing concern (De Freitas et al., 2023; Schepman & Rodway, 2020; Stein et al., 2024). One popular scale to measure attitudes towards AI is the ATTARI, consisting of twelve items in its original form (Stein et al., 2024). Across two studies encompassing three large samples, we introduce the ATTARI-6 and the ATTARI-3, two short forms of the ATTARI scale. The ATTARI-6 is a concise version that incorporates both positively and negatively worded items, whereas the ATTARI-3 is intended for surveys requiring an ultra-brief measure.

## Measuring Attitudes Toward AI and the ATTARI-12

Understanding how individuals evaluate AI requires psychometrically sound instruments that capture the multifaceted nature of attitudes. Attitudes are traditionally conceptualized as evaluations comprising cognitive, affective, and behavioral components (Rosenberg & Hovland, 1960; Zanna & Rempel, 2008). Yet, research on AI attitudes has often relied on ad hoc items or narrow instruments that focus on specific applications (e.g., self-driving cars, service robots). Although such instruments are valuable, they limit

comparability across studies and neglect the broader evaluation of AI as a technological phenomenon that transcends specific use cases (Stein et al., 2024).

In recent years, several dedicated scales on more general constructs reflecting attitudes towards AI have been developed, including the General Attitudes Toward Artificial Intelligence Scale (GAAIS; Schepman & Rodway, 2020), the Attitudes Toward Artificial Intelligence Scale (ATAI; Sindermann et al., 2021), the AI Attitude Scale (AIAS; Grassini, 2023), the AI Anxiety Scale (AIAS; Wang & Wang, 2022), and the Threats of Artificial Intelligence Scale (TAI; Kieslich et al., 2021). Whereas these measures clearly contributed to the field, they also contained notable limitations (Kucherenko, 2025). Two of the named instruments focused exclusively on negative evaluations (AIAS, Wang & Wang, 2022, TAI, Kieslich et al., 2021), whereas one scale consists of positive item wordings only in its final iteration (AIAS, Grassini, 2023). Two other scales imposed multifactorial structures, yielding separate scores for positive attitudes and negative attitudes (GAAIS, Schepman & Rodway, 2020) or for acceptance and fear (ATAI, Sindermann et al., 2021), respectively. Although a multidimensional measurement approach may be beneficial in some circumstances, other researchers prefer a one-dimensional scale that is easy to interpret. Importantly, at the time of development of the ATTARI scale, no scale had incorporated the full tripartite model of attitudes (cognition, affect, behavior), despite its centrality in psychological theory (Ajzen, 2001).[1]

Against this background, Stein et al. (2024) introduced the Attitudes Toward Artificial Intelligence scale (ATTARI-12). The ATTARI-12 is a 12-item, unidimensional measure that integrates the cognitive, affective, and behavioral facets of attitudes toward AI. Its items are

[1] In the educational realm, scale developments incorporated the tripartite model to assess AI attitudes (Katsantonis & Katsantonis, 2024; Suh & Ahn, 2022). Both include a substantial number of items, however. The *Student Attitude Toward Artificial Intelligence- Scale* (SATAI, Suh & Ahn, 2022) consists of 26 items. The Greek adapted version of the SATAI (Katsantonis & Katsantonis, 2024) consists of 25 items.

balanced between positively and negatively worded statements, thus reducing acquiescence bias. Across multiple samples in the United States and Germany, the ATTARI-12 demonstrated excellent internal consistency (Cronbach's α around or above .90). It has been translated to several languages and psychometrically tested versions are available in Chinese (P. Li et al., 2026), Japanese (Fujikawa et al., 2026), Korean (Lee et al., 2025), Polish (Juranek & Pochwatko, 2025), and Romanian (Maier et al., 2025). A related scale was developed to assess domain-specific attitudes towards AI in the fields of work, health, and education (ATTARI-WHE, Gnambs et al., 2025). Convergent validity of the ATTARI-12's assessment of overarching AI-related attitudes with more specific attitudes toward robots and voice assistants was established, whereas it was found to not be affected by social desirability (Stein et al., 2024). The scale thus provides a theoretically grounded, psychometrically robust, and context-independent instrument for assessing general attitudes toward AI. The ATTARI-12 is highly cited and regularly demonstrates strong psychometric properties (e.g., Montag et al., 2026).

Beyond establishing the scale's psychometric quality, Stein et al. (2024) also examined predictors of AI attitudes. They found that agreeableness predicted more positive views of AI, whereas conspiracy mentality predicted more negative views. Age also emerged as a predictor, with older individuals holding more negative attitudes. These findings highlight the importance of both dispositional and demographic variables in shaping public opinion about AI and underscore the need for standardized, validated measures to advance cumulative knowledge.

**The Case for ATTARI-Short Scales**

Since its introduction, the ATTARI-12 has been employed by many research teams worldwide. Still, a 12-item measure may present challenges in large-scale surveys, field studies, or applied contexts where time is limited (Rolstad et al., 2011). Shorter scales offer advantages in terms of efficiency, reduced respondent burden, and broader applicability. For

example, widely used short forms such as the Ten Item Personality Inventory (Gosling et al., 2003) and the PHQ-2 (Kroenke et al., 2003) have enabled rapid yet valid assessments of personality and depression, respectively.

The development of the ATTARI short forms was motivated by a lack of a shorter, unidimensional scale that provided one score for a global attitude towards AI, and that at the same time reflected the tripartite structure of attitudes (for additional information on the currently available approaches to measuring attitudes towards AI, see e.g., Kucherenko, 2025; Montag et al., 2026; Schepman & Rodway, 2025). A challenge for shorter scales, however, is low reliability. In the field of attitudes towards AI, the ATAI instrument (Sindermann et al., 2021) consists of two dimensions that are represented by two and three items, respectively. The desired brevity, however, comes with less-than-ideal reliability scores. The internal consistencies of the ATAI acceptance and the ATAI fear scales ranged around .64 and .66 for the English and German language versions (Sindermann et al., 2021).

Thus, scale shortening must be conducted with care to preserve reliability and validity (Smith et al., 2000; Ziegler et al., 2014). When guided by both theoretical reasoning and empirical evidence, abbreviated versions can retain the conceptual breadth of the original instrument while increasing usability. The ATTARI-12's unidimensional structure and strong item loadings make it well-suited for the development of abbreviated versions. Shorter forms would enable the inclusion of AI attitudes in multidisciplinary surveys, experimental studies, and organizational assessments where survey space is constrained. Moreover, ultra-brief versions could facilitate quick assessments of public sentiment in policy, education, or industry contexts.

**Item Selection and Study Overview**

The present work introduces and validates two short forms of the ATTARI-12: a 6-item version (ATTARI-6) and a 3-item version (ATTARI-3). All five authors of the present paper took part in focus discussions to determine the ideal item selection for the newly

developed short forms. Our goal was to preserve coverage of positive and negative wordings in the six-item version and to retain the assessment of cognitive, affective, and behavioral facets in both short versions. We further wished to minimize semantic redundancy and potential semantic imprecisions of the items by selecting the best items for the short versions. In terms of statistical indicators, selection criteria were factor loadings and item popularity, as obtained in a prior study with an English-language sample (Stein et al., 2024, Study 1). Table 1 shows the respective results.

The resulting ATTARI-6 is designed as a moderately shortened measure balancing efficiency with comprehensiveness. It includes one positively and one negatively formulated item for each of the three attitude facets (cognitive, affective, behavioral). The ATTARI-3 provides an ultra-brief alternative for rapid or large-scale assessments. It consists of one positively worded item for each of the three attitude facets. Although our approach was meant to preserve the semantic breadth of the original scale, we need to acknowledge that shortened scales always reduce the amount of information gathered.

The psychometric assessment of the ATTARI short forms was guided by three primary aims. First, we tested the factorial validity of the ATTARI-6 and ATTARI-3, evaluating their alignment with the unidimensional structure of the ATTARI-12 and, second, we assessed the reliability of the short forms. Third, we examined construct validity by exploring associations with external variables previously linked to AI attitudes with a focus on usage of generative AI and personality traits (Stein et al., 2024).

In the sections that follow, we detail the measurement model, the validation, and comparative analysis of the ATTARI-6 and ATTARI-3. These studies were conducted with the English (Studies 1a and Study 2) and German (Study 1b) ATTARI versions. We conclude by discussing implications for theory, future research, and applied practice in the rapidly evolving landscape of AI technologies.

## Study 1: Measurement Structure of the ATTARI-6 and ATTARI-3

### Materials and Method

#### *Participants and Procedure*

Two independent samples were recruited from the online research panel *Prolific*. The first sample was limited to residents of the United States and the United Kingdom (Study 1a), whereas the second sample comprised participants located in Germany (Study 1b). Eligibility criteria for both samples required participants to be at least 18 years of age and fluent in English (Study 1a) or German (Study 1b). The data underlying Study 1a were retrieved from a larger study with a study goal that was unrelated to AI attitudes (i.e., the ATTARI served as a filler scale).

Participants who completed the survey received a remuneration of 1.50 GBP (Study 1a) or 0.60 GBP (Study 1b). Data quality was ensured by excluding non-diligent respondents based on extremely short completion times or failing one or more attention checks (see Online Supplement S1). The final samples included 1,023 participants (448 women, 571 men, 4 other) in Study 1a and 994 participants (382 women, 599 men, 13 other) in Study 1b. The mean age of respondents in Study 1a was 39.36 years ($SD$ = 12.97, $Min$ = 18, $Max$ = 90), whereas the mean age in Study 1b was 32.4 years ($SD$ = 10.40, $Min$ = 18, $Max$ = 74). Both samples were relatively highly educated, with approximately 70% (Study 1a) and 56% (Study 1b) reporting a completed post-secondary education (e.g., undergraduate, bachelor's, or equivalent degree).

#### *Ethics Statement*

Protocols for Studies 1a and 1b were approved by the responsible ethics committee at the University of Würzburg, Germany. All participants provided informed consent prior to participation, and their rights, autonomy, and confidentiality were upheld throughout the research process.

### *Transparency and Openness*

The analyses were conducted in *R* (R Core Team, 2025) using the *lavaan* (Version 0.6-20; Rosseel, 2012) and *semTools* packages (Jorgensen et al., 2025). The data and analysis code are provided at https://osf.io/byq6n.

### *Measure*

Studies 1a and 1b both adhered to the same procedure. Following an introductory definition of the AI concept (see Appendix), participants completed the twelve items of the ATTARI-12 (Stein et al., 2024). All items had to be rated on five-point response scales ranging from 1 (*strongly disagree*) to 5 (*strongly agree*). Afterwards, questions on core demographic indicators were asked.

### *Statistical Analyses*

The reliability of the short ATTARI forms was assessed using coefficient alpha and categorical omega (Flora, 2020). To examine the internal structure of the ATTARI-6, we conducted confirmatory factor analyses (CFAs) for ordinal data using diagonally weighted least square estimation (C. Li, 2016).

Following Stein et al. (2024), three theory-driven CFA models were specified for each language version. Model 1 posited a unidimensional structure with all six items loading on a single latent factor. Model 2 specified a bifactor structure comprising a general factor for all six items and two orthogonal facet factors representing affective and behavioral attitudes. Following Eid and colleagues (2017), no specific factor was specified for the cognitive facet which therefore served as the reference facet. Model 3 examined potential wording effects by specifying an orthogonal method factor for negatively worded items in addition to a general factor for all six items. Graphical representations for all models are provided in Figure S2.1 of the supplementary material. The latent factors in these models were identified by fixing the factor variances to 1 and the factor means to 0.

Model fit was evaluated using the standardized root mean square residual (SRMR), the comparative fit index (CFI), and the Tucker-Lewis index (TLI)[2]. Following Schermelleh-Engel et al. (2003), values of SRMR ≤ .05/.10 and CFI / TLI ≥ .97/.95 were interpreted as good/acceptable model fit.

Measurement invariance across language versions (English versus German) was examined using multi-group CFAs (Wu & Eastbrook, 2016). Threshold and loading invariance were tested by imposing equality constraints on item thresholds and factor loadings, respectively. Nested model comparisons were evaluated using the chi-squared difference test and changes in CFI, with a statistically significant test ($p < .05$) combined with a decrease in CFI greater than .01 interpreted as evidence of non-invariance (Svetina et al., 2020). To assess the practical relevance of detected non-invariance, standardized mean differences in expected item scores ($d_{MACS}$; Nye et al. 2019) were calculated. Following Gunn et al. (2020), $d_{MACS}$ values below .20 were considered negligible, values between .20 and .40 as moderate, and values above .40 as large and potentially problematic for fair group comparisons.

## Results and Discussion

### *Descriptive Statistics for the ATTARI-6 and ATTARI-3*

As shown in Table 2, the item means and standard deviations for the ATTARI-6 and ATTARI-3 showed no floor or ceiling effects in either sample. All items exhibited substantive variability across respondents in both the English language and German samples. Item discrimination indices were uniformly high. For the ATTARI-6, corrected item-total correlations ranged from .71 to .83 in Study 1a and were only slightly lower in Study 1b (*Min* = .71, *Max* = .77). The items included in the shorter ATTARI-3 also demonstrated strong discrimination, with coefficients around or above .70 in both samples.

[2] We refrain from using the root mean squared error of approximation because it has an excessive Type I error rate in models with small degrees of freedom when applying conventional cutoffs for model evaluation and tends to reject even properly specified models (Kenny et al., 2015).

At the scale level, the mean score of the ATTARI-6 and ATTARI-3 closely resembled those of the original ATTARI-12 in both samples, while their variances were slightly larger (see Table 3). Scale scores for both the long and brief versions were approximately normally distributed, as indicated by the skewness and kurtosis values in Table 3 and by the violin plots presented in the supplementary materials (Figure S3.1). Finally, the ATTARI-6 and ATTARI-3 scores were highly correlated with the corresponding ATTARI-12 scores in both samples, $r = .98$ ($p < .001$) and $r = .95$ ($p < .001$), indicating a high degree of overlap between the brief versions and the original scale.

### *Factor Structure of the ATTARI-6*

The dimensionality of the ATTARI-6 was examined by comparing three competing confirmatory factor models. Model fit indices for all models are reported in Table 4. The results indicated that the unidimensional model which specified a single latent factor for the six items (Model 1) demonstrated excellent fit in both samples. Specifically, the SRMR values were .04 (English sample) and .03 (German sample), TLI values were .95 and .97, and CFI values were .98 and .99, respectively. The standardized loadings in Model 1 were substantial in both samples (see Table 2). In the English sample, loadings ranged from .71 to .83 (*Mdn* = .73), whereas in the German sample they ranged from .64 to .77 (*Mdn* = .70).

Comparisons between the competing CFA models provided no evidence for multidimensionality introduced by the attitude facets because the respective bifactor model (Model 2) did not improve model fit relative to the unidimensional model (see Table 4). In contrast, Model 3, which modeled wording effects for negatively keyed items, yielded a statistically superior fit in both samples. However, acknowledging the method factor had only a marginal impact on the standardized loadings of the general factor (see Table 2). For negatively worded items, median loadings decreased by only *Mdn* = .06 in the English sample and *Mdn* = .04 in the German sample. Accordingly, the general factor accounted for approximately 74% or more common item variance in the English sample and at least 84% in

the German sample. Taken together, these findings indicate that the ATTARI-6 exhibits an essentially unidimensional measurement structure, which is only negligibly affected by wording effects.

Systematic dimensionality analyses were not conducted for the ATTARI-3 because a latent factor model with three indicators is just-identified and, therefore, does not inform about model fit. Nonetheless, the standardized factor loadings reported in Table 2 indicate a well-defined latent factor in both samples, with loadings ranging from .79 to .90.

***Reliability of the ATTARI-6 and ATTARI-3***

Table 3 summarizes the internal consistency estimates for the brief ATTARI versions and, for comparison, the original ATTARI-12. Despite reducing the length of the instrument by one half (ATTARI-6) or three quarters (ATTARI-3), reliability remained high in both samples. In the English sample, categorical omega coefficients were .92 for the ATTARI-6 and .89 for the ATTARI-3, only marginally lower than the reliability of the ATTARI-12 ($\omega$ = .96). Comparable results were observed in the German sample, with excellent omega reliabilities of .89 for the ATTARI-6 and .83 for the ATTARI-3. Accounting for the method factor in the ATTARI-6 hardly affected the reliability of the general factor. Estimates of omega hierarchical remained high ($\omega_h$ = .87 in the English sample and $\omega_h$ = .89 in the German sample), indicating that the general factor remained highly reliable even when variance attributable to item wording was partialled out.

***Measurement Invariance of the ATTARI-6 and ATTARI-3***

Measurement invariance across language versions (Study 1a: English versus Study 1b: German) was examined using multi-group CFAs. The results of the respective analyses supported strong measurement invariance for both brief ATTARI versions. For the ATTARI-6, constraining item thresholds to equity across language groups resulted in a statistically significant deterioration in model fit relative to the configural model, $\Delta\chi^2(df = 12) = 23.53$, $p$ = .024. However, global goodness-of-fit indices were virtually unchanged and indicated

excellent fit for the configural model (CFI = .98, TLI = .97, SRMR = .04) and the threshold-invariance model (CFI = .99, TLI = .99, SRMR = .04). Similarly, imposing equality constraints on both item thresholds and factor loadings led to a significant difference test, $\Delta\chi^2(df = 5) = 32.27$, $p < .001$, but had negligible impact on approximate fit indices (CFI = .99, TLI = .99, SRMR = .04). Moreover, the standardized mean differences in expected item scores ($d_{MACS}$) indicated small and largely negligible differences between language versions, with values ranging from –0.21 (Item 3) and 0.37 (Item 11).

Measurement invariance across English and German language versions was also supported for the ATTARI-3. Although difference tests indicated statistically significant deteriorations in fit for models constraining the item thresholds, $\Delta\chi^2(df = 6) = 12.82$, $p = .046$, or both thresholds and loadings, $\Delta\chi^2(df = 2) = 12.56$, $p = .002$, relative to the configural model, the approximate goodness-of-fit indices remained unaffected. Specifically, CFI and TLI values were 1.00 and SRMR values fell below .01 in all constrained models. In line with these results, the $d_{MACS}$ values were small (–0.21 to 0.37) and did not indicate substantive item-level non-invariance.

***Summary of Studies 1a and 1b***

Across two large samples, the English and German language versions of the ATTARI-6 and ATTARI-3 demonstrated strong psychometric properties. Both short forms maintained high reliability despite substantial reductions in length and produced score distributions comparable to the original ATTARI-12. The ATTARI-6 further exhibited an essentially unidimensional measurement structure, with little evidence of meaningful multidimensionality. Finally, both brief versions exhibited strong measurement invariance across language versions which supports their use in multilingual research contexts. We wish to add that the very high correlations between the short forms and ATTARI-12 are partly attributable to overlapping items. Thus, these correlations should not be interpreted as

independent evidence of validity. The notion of validity was the main focus of our second study.

## Study 2: Validity of the ATTARI Short Forms

Two aims guided Study 2. First, we wanted to test whether the reliability estimates and the factor structure of the ATTARI short forms replicated in an independent sample. Second, we examined the construct validity of the ATTARI short forms by comparing associations between the three forms and theoretically and empirically related variables.

The first variable that served as a validity indicator was AI anxiety (Wang & Wang, 2022). Given that anxiety towards the attitude object — in our case: AI — should be negatively related to the attitude towards this attitude object (e.g., Rosenberg & Hovland, 1960), we expected that more positive attitudes towards AI (as measured with any of the three ATTARI forms) would be associated with lower AI anxiety. Specifically, we chose the AI anxiety subscale *job replacement* (Wang & Wang, 2022) as a relevant indicator for this investigation. Next, we expected that more positive attitudes towards AI (as measured with any of the three ATTARI forms) would be related to more frequent usage and more positive usage intentions in terms of actual daily AI usage in private contexts and preferences for using AI at work.

In their series of studies, Stein and colleagues (2024) showed that more positive attitudes towards AI (as measured with the original ATTARI-12 long form) were associated with higher agreeableness, as well as with lower conspiracy mentality. We therefore expected that all three ATTARI forms were similarly associated with these two individual difference variables. Finally, theory and research suggest that humans are particularly wary when it comes to computers or robots that are described to have mind, particularly digital entities with feelings (e.g., Appel et al., 2020; Gray & Wegner, 2012; Stein & Ohler, 2017). Hence, we hypothesized that positive attitudes towards AI (as measured with any of the three ATTARI

forms) would be associated with less eeriness in response to a robot that can feel. Further variables of interest were age and gender, for which no directed hypotheses were formulated.

Importantly, our goal was to compare the expected associations when administering the ATTARI-6 or the ATTARI-3, as compared to the ATTARI-12. To this end, Study 2 followed an experimental design; participants were randomly assigned to one of the three ATTARI versions. In addition to descriptive comparisons of the respective associations, we conducted non-inferiority tests of the short forms as compared to the long version (Walker, 2019).

## Method

### *Participants*

We aimed for a US sample that reflected population quota in terms of education, age, and ethnic background using the platform *Prolific*. Sample size was determined a priori. Based on the planned non-inferiority tests, 350 participants per condition were required, leading to a total aspired sample size of 1050. To account for potential exclusions, we recruited 1,238 US residents with English as their first language. Individuals who had completed Study 1 were not eligible for participation in this study. Participants were compensated with 0.75 GBP. Of our initial sample, 84 participants had to be excluded because they did not pass one or more of the pre-registered inclusion criteria. They were either flagged by our survey software's (*Qualtrics*) bot detection tool ($n$ = 19), answered incorrectly on an instructed response item ($n$ = 22), gave meaningless responses to an open-ended question on the study's topic ($n$ = 34), or showed a total response time that was too short to process the items ($n$ = 17; Huang et al., 2012). The final sample consisted of 1,154 participants (50.2% women, 48.2% men, 1.6% other) who were between 18 and 87 years old ($M$ = 46.47, $SD$ = 16.62), with 64% self-identifying as White, 11.8% as Black, 10.9% as Mixed, 6.0% as Asian, and 7.8% as Other.

### *Ethics Statement*

The protocol for Study 2 was approved by the responsible ethics committee at the University of Würzburg, Germany. All participants provided informed consent prior to participation, and their rights, autonomy, and confidentiality were upheld throughout the research process.

### *Transparency and Openness*

The analyses were conducted in *R* (R Core Team, 2025) using the *lavaan* (Version 0.6-20; Rosseel, 2012) and *semTools* packages (Jorgensen et al., 2025). The data and analysis code are provided at https://osf.io/byq6n. Study 2 was preregistered at https://aspredicted.org/tx23c2.pdf.

### *Measures*

**ATTARI scales.** Participants were randomly assigned to complete the ATTARI-12, the ATTARI-6, or the ATTARI-3. The items had to be rated on five-point response scales ranging from 1 (*strongly disagree*) to 5 (*strongly agree*). Reliability estimates are provided in the Results section.

**Agreeableness.** Agreeableness was assessed with the items of the Big Five Inventory (John et al., 2008). The scale included nine items, e.g., “I am someone who is considerate and kind to almost everyone”. Five-point response scales ranging from 1 (*disagree strongly*) to 5 (*agree strongly*) were provided (Cronbach’s $\alpha = .84$).

**Conspiracy Mentality.** We used the Conspiracy Mentality Questionnaire (Bruder et al., 2013), a measure of people’s general susceptibility to conspiracy theories and beliefs. It consists of five items, e.g., “I think that events which superficially seem to lack a connection are often the result of secret activities”. Scale anchors were 1 (*disagree strongly*) and 5 (*agree strongly*, Cronbach’s $\alpha = .87$)

**AI-Related Job Preferences.** The scale consisted of four items that were adapted from prior research (Stein et al., 2024). The item wordings were applicable for individuals

who were currently in or out of the workforce, e.g., “Regardless of whether you are currently employed, what are your preferences regarding AI in the workplace? – I prefer a job in which I engage with AI constantly”. Five-point response scales were provided with a scale ranging from 1 (*disagree strongly*) to 5 (*agree strongly*, Cronbach’s α = .90, see Supplement S4 for all item wordings).

**AI Anxiety (Job Replacement).** AI anxiety was measured with the subscale *job replacement* of the AI Anxiety Questionnaire by Wang and Wang (2022), which consists of six items e.g., “I am afraid that widespread use of humanoid robots will take jobs away from people”. Responses were recorded on a five-point scale ranging from *not at all* (1) to *extremely* (5), Cronbach’s α = .92.

**Eeriness in Response to a Robot with Mind.** A textual vignette from a study by Appel and colleagues (2020) was presented to introduce participants to a (fictitious) robot with needs and the ability to feel emotions (experiencer robot “Ellix”). After this brief description, participants indicated their feelings of eeriness with three items from Gray and Wegner (2012), e.g., “Thinking about robot ‘Ellix’ makes me feel uneasy”. The items went with a five-point scale ranging from *not at all* (1) to *extremely* (5), Cronbach’s α = .96.

### *Statistical Analyses*

In the first step, the reliability and dimensionality of the brief ATTARI versions were evaluated using the same procedures as in Study 1. Internal consistency was assessed using coefficient alpha and categorical omega (Flora, 2020); the factor structure was examined with CFA for ordinal data (C. Li, 2016).

In the second step, correlations between the ATTARI-12 and validity indicators were compared with those obtained for shorter ATTARI versions using a non-inferiority framework (Walker, 2019). The aim was to determine whether the correlations obtained with the short forms were not meaningfully smaller than those for the full scale. A non-inferiority margin of .20 was specified a priori, such that significant results indicate that the correlation

for a short version does not fall more than .20 below that of the ATTARI-12. This margin was selected based on empirical distributions of correlations and reflects the typical magnitude of correlations observed in psychological research (Gignac & Szodorai, 2016; Weinerová et al., 2022). Inference was based on a nonparametric bootstrap procedure with stratified resampling. A total of 5,000 bootstrap samples was drawn. Confidence intervals for the difference in correlations were estimated using the bias-corrected and accelerated method, which adjusts for bias and skewness in the bootstrap distribution. Non-inferiority was evaluated using a one-sided criterion: specifically, a bootstrap-based *p*-value was computed as the proportion of bootstrap samples in which the observed difference fell below the non-inferiority margin.

## Results and Discussion

### *Reliability and Factor Structure of the ATTARI-6 and ATTARI-3*

Both short forms demonstrated high internal consistency. For the ATTARI-6, categorical omega and coefficient alpha were .96 and .95, respectively, whereas both indices were .93 for the ATTARI-3. Thus, the reliabilities of the short forms were only marginally lower than those observed of the full ATTARI-12 (α = .96, ω = .98). As shown in Table 5, all items exhibited strong corrected item–total correlations that ranged from .80 to .88 for the ATTARI-6 and from .84 to .88 for the ATTARI-3. Violin plots of the observed distributions are shown in Table S3.2 of the supplementary material.

Confirmatory factor analyses for the ATTARI-6 replicated the findings from Study 1. A unidimensional model with a single latent factor provided an excellent fit to the data, $\chi^2$ (*df* = 9) = 139.45, SRMR = .03, TLI = .99, CFI = .99. Extending the model to a bifactor specification to account for potential method effects results in only negligible improvements in model fit, $\chi^2$ (*df* = 6) = 98.23, SRMR = .02, TLI = .98, CFI = .99. Standardized factor loadings for both short versions were uniformly high (see Table 5), ranging from .85 to .94 (*Mdn* = .90) for the ATTARI-6 and from .90 to .97 (*Mdn* = .93) for the ATTARI-3. Overall,

these findings support the unidimensional structure and high measurement precision of both ATTARI short forms.

### *Correlations with Validity Indicators*

Correlations between the ATTARI versions and external validity indicators are reported in Table 6. As expected, AI attitudes were strongly associated with career interests in AI-related professions and with AI anxiety concerning job replacement. In addition, ATTARI scores showed significant associations with agreeableness and with feelings of eeriness in response to a humanoid robot. In contrast, correlations with conspiracy mentality and age turned out to be small.

Importantly, the pattern of correlations was highly similar across ATTARI versions. Absolute differences in correlations between the full ATTARI-12 and their short forms ranged from .00 to .20. The significant non-inferiority tests further indicated that the correlations obtained with the ATTARI-6 and ATTARI-3 were not meaningfully smaller than those observed for the full scale. Taken together, these findings suggest that shortening the instrument did not substantially alter its nomological network, thus supporting the validity of the ATTARI short forms.

## General Discussion

In a technological era shaped by powerful AI systems and related innovations, it has become crucial to understand how individuals assess the notion of artificially intelligent technology in the first place—in order to ensure societally responsible transformation processes. Scholars have developed several measurement tools that capture people's AI-related attitudes from a variety of perspectives (e.g., domain-specific vs. more general assessments; positive views vs. negative views vs. both). The ATTARI-12 questionnaire (Stein et al., 2024) is among the most well-received contributions in this regard. It offers a way to measure individuals' general attitude toward AI as a single composite score, independent of specific use cases or fields of application, yet, reflecting the tripartite structure

of attitudes. Given its strong psychological foundation and empirical validation, the scale has been highly popular. Translations of the original English and German scales are available in several languages.

The short forms introduced in the present work preserve the original structure of the ATTARI-12 by incorporating cognitive, affective, and behavioral aspects of attitudes towards AI. In addition, the ATTARI-6 consists of positively and negatively worded items. With only half (ATTARI-6) or one fourth (ATTARI-3) of the original items, the short forms demonstrated strong psychometric quality. In all samples, reliability estimates ranged around .90 for the ATTARI-6 and exceeded .80 for the ATTARI-3. Tests of factorial validity and measurement invariance corroborate the soundness of assessing a global attitude score with both measures. At the same time, the short forms showed associations with criteria of interest that were similar in size to the associations based on the long scale (as indicated by significant non-inferiority tests).

The ATTARI-6 and ATTARI-3 scales can be applied whenever research constraints call for short versions. This is particularly the case for large-scale surveys that often aim for a population representative assessment. Given that attitudes towards AI will be one among many variables of interest in such settings, scale brevity is of particular priority. Such large-scale surveys are often used to track certain population parameters over time. For example, Gnambs and Appel (2019) analyzed data from the Eurobarometer surveys in the years 2012, 2014, and 2017. In each year, the survey contained the same three items on attitudes towards *robots*. The authors found that the attitudes towards robots (general appraisals) declined and delineated intriguing between-country differences. Given the impact of AI on our societies, we believe that the ATTARI-3 will be ideal to be included in large-scale surveys worldwide. Other researchers may welcome the brevity of the shorter versions but consider the combination of positively and negatively worded items (forming a balanced scale, like the ATTARI-6) to be indispensable. A benefit of such balanced scales is that reversed items

activate participants to think more deeply about the item content rather than responding mindlessly, thereby increasing validity (e.g., Weijters & Baumgartner, 2022). Negatively worded items further reduce the potential impact of the well-known acquiescence bias: Some participants tend to agree to statements, largely irrespective of the item content (e.g., Cronbach, 1960). Acquiescence bias has been observed in many fields of the social sciences, and balanced scales (scales with positively and negatively worded items) are considered one way to reduce the systematic impact of this bias (e.g., Kuru & Pasek, 2016).

By offering two new validated versions of the ATTARI scale, our work provides flexibility and enhances the methodological toolkit for studying attitudes toward AI. After our studies were planned and conducted, we became aware of a short form of a main alternative instrument to measure attitude towards AI, the GAAIS (GAAIS-10, Schepman & Rodway, 2026). This instrument comprises two subscales (positive, negative) that consist of 5 items each. Researchers who favor a separate assessment of positive and negative attitudes towards AI may refer to this alternative new scale.

**Limitations and Future Research Directions**

Arguably, AI-related developments—and the transformations they invoke across numerous areas of daily life—happen at a staggering pace. Against this backdrop, we made sure to specifically build our novel ATTARI short forms around those items from the original scale that promise the most robust application for years to come (e.g., very clear-cut descriptive statements such as “I want to use technologies that rely on AI.”). Given the item wordings, we are confident that the ATTARI-6 and ATTARI-3 scales will be robust to changes regarding how AI technologies look like or develop over the years to come. Nevertheless, it seems worthwhile to continuously monitor how potential advancements and societal changes affect the applicability of the ATTARI inventories in different contexts.

Relatedly, future research is encouraged to examine the trajectory of attitudes towards AI. This could involve international perspectives, as attitudes towards AI may differ

meaningfully across countries. The research presented here gave insights into the psychometric properties of the English and German language versions of the ATTARI-6 and the ATTARI-3. We encourage researchers to examine the ATTARI short forms in other language versions.

The ATTARI is an instrument that measures global attitudes towards AI. Researchers who are interested in a specific field of application might find these global measures to be too remote from this specific field of application. Under such circumstances other instruments may provide a better fit. Based on the ATTARI, Gnambs and colleagues introduced the ATTARI-WHE (Gnambs et al., 2025) consisting of items that cover the fields of work, health, and education.

Further methodological caveats require attention: By design, each attitude component of the ATTARI-3 is represented by only a single item, making the measure suitable as an ultra-brief indicator but limiting the assessment of the tripartite attitude structure. Accordingly, systematic analyses of dimensionality for the ATTARI-3 were not possible. Moreover, the evidence on the nomological validity presented in Study 2 is largely based on cross-sectional correlations with other self-report scales. The testing of scale validity is never a completed endeavor (e.g., Messick, 1995); every new study contributes to the assessment of a scale's validity. Thus, we particularly encourage future research on behavioral corollaries of the ATTARI short forms.

We further note that all samples in this set of studies were online samples, and the participants were members of the Prolific panel. Although we have no reason to doubt the reliability and validity of the scales in other contexts (e.g., paper-and-pencil surveys, phone surveys) or with other samples (e.g., student-only samples, individuals with decidedly low technological affinity), future research should provide additional insights into the generalizability of our results.

Finally, we assume that attitudes towards AI are relatively stable, with scores reflecting a large trait component. That said, like all attitude measures, the ATTARI short forms introduced here may include a state component as well. Future research is encouraged to quantify the extent of trait and state variance. Experimental variation may result in intraindividual variance that could be analyzed through latent-state-trait analyses (e.g., Steyer et al., 1999).

**Conclusion**

Measures of general attitudes toward AI are increasingly used in surveys and experiments across psychology and related fields. The ATTARI-6 and the ATTARI-3 were introduced as two short forms of the original ATTARI scale (Stein et al., 2024) that retain the tripartite structure of attitudes. The short forms consist of six and three items, respectively, and across three samples they showed excellent psychometric properties in the German and English versions. We believe that both short forms will be used extensively in future research.

**References**

Ajzen, I. (2001). Nature and operation of attitudes. *Annual Review of Psychology, 52*, 27–58. https://doi.org/10.1146/annurev.psych.52.1.27

Appel, M., Izydorczyk, D., Weber, S., Mara, M., & Lischetzke, T. (2020). The uncanny of mind in a machine: Humanoid robots as tools, agents, and experiencers. *Computers in Human Behavior, 102*, 274–286. https://doi.org/10.1016/j.chb.2019.07.031

Bruder, M., Haffke, P., Neave, N., Nouripanah, N., & Imhoff, R. (2013). Measuring individual differences in generic beliefs in conspiracy theories across cultures: Conspiracy Mentality Questionnaire. *Frontiers in Psychology, 4*, 225. https://doi.org/10.3389/fpsyg.2013.00225

Eid, M., Geiser, C., Koch, T. & Heene, M. (2017). Anomalous results in G-factor models: Explanations and alternatives. *Psychological Methods, 22*, 541–562. https://doi.org/10.1037/met0000083

Cronbach, L. J. (1960). *Essentials of psychological testing* (2nd ed.). Harper & Row.

De Freitas, J., Agarwal, S., Schmitt, B., & Haslam, N. (2023). Psychological factors underlying attitudes toward AI tools. *Nature Human Behaviour, 7*(11), 1845–1854. https://doi.org/10.1038/s41562-023-01734-2

Flora, D. B. (2020). Your coefficient alpha is probably wrong, but which coefficient omega is right? A tutorial on using R to obtain better reliability estimates. *Advances in Methods and Practices in Psychological Science, 3*(4), 484-501. https://doi.org/10.1177/2515245920951747

Fujikawa, H., Mori, H., Kondo, K., Nishizaki, Y., Yano, Y., & Naito, T. (2026). Adaptation of the Japanese Version of the 12-item Attitudes Towards Artificial Intelligence Scale for medical trainees: Multicenter development and validation study. *JMIR Medical Education*, *12*, Article e81986. https://doi.org/10.2196/81986

Gignac, G. E., & Szodorai, E. T. (2016). Effect size guidelines for individual differences researchers. *Personality and Individual Differences, 102*, 74-78. https://doi.org/10.1016/j.paid.2016.06.069

Gnambs, T., & Appel, M. (2019). Are robots becoming unpopular? Changes in attitudes towards autonomous robotic systems in Europe. *Computers in Human Behavior, 93*, 53-61. https://doi.org/10.1016/j.chb.2018.11.045

Gnambs, T., Stein, J. P., Appel, M., Griese, F., & Zinn, S. (2025). An economical measure of attitudes towards artificial intelligence in work, healthcare, and education (ATTARI-WHE). *Computers in Human Behavior: Artificial Humans*, *3*, Article 100106. https://doi.org/10.1016/j.chbah.2024.100106

Gosling, S. D., Rentfrow, P. J., & Swann, W. B., Jr. (2003). A very brief measure of the Big-Five personality domains. *Journal of Research in Personality, 37*(6), 504–528. https://doi.org/10.1016/S0092-6566(03)00046-1

Grassini, S. (2023). Development and Validation of the AI Attitude Scale (AIAS-4): A Brief Measure of Attitude Toward Artificial Intelligence, *Frontiers in Psychology, 14*, Article 1191628. https://doi.org/10.3389/fpsyg.2023.1191628

Gray, K., & Wegner, D. M. (2012). Feeling robots and human zombies: Mind perception and the uncanny valley. *Cognition, 125*(1), 125–130. https://doi.org/10.1016/j.cognition.2012.06.007

Gunn, H. J., Grimm, K. J., & Edwards, M. C. (2020). Evaluation of six effect size measures of measurement non-invariance for continuous outcomes. *Structural Equation Modeling*, *27*(4), 503-514. https://doi.org/10.1080/10705511.2019.1689507

Huang, J. L., Curran, P. G., Keeney, J., Poposki, E. M., & DeShon, R. P. (2012). Detecting and deterring insufficient effort responding to surveys. *Journal of Business and Psychology, 27*(1), 99–114. https://doi.org/10.1007/s10869-011-9231-8

John, O. P., Naumann, L. P., & Soto, C. J. (2008). Paradigm shift to the integrative big-five trait taxonomy: History, measurement, and conceptual issues. In O. P. John, R. W. Robins, & L. A. Pervin (Eds.), *Handbook of personality: Theory and research* (pp. 114-158). New York, NY: Guilford Press.

Jorgensen, T. D., Pornprasertmanit, S., Schoemann, A. M., & Rosseel, Y. (2025). *semTools*: *Useful Tools for Structural Equation Modeling* (Version 0.5-7) [Computer software]. https://doi.org/10.32614/CRAN.package.semTools

Juranek, J., & Pochwatko, G. (2025). Measurement of general attitude toward AI and its relationship with AI anxiety, attitude toward robots and belief in human nature uniqueness: Polish adaptation of the ATTARI-12 questionnaire. *Polish Psychological Bulletin, 57*, 88-99. https://doi.org/10.24425/ppb.2026.153995

Katsantonis, A., & Katsantonis, I. G. (2024). University students' attitudes toward artificial intelligence: An exploratory study of the cognitive, emotional, and behavioural dimensions of AI attitudes. *Education Sciences*, *14*(9), Article 988. https://doi.org/10.3390/educsci14090988

Kenny, D. A., Kaniskan, B., & McCoach, D. B. (2015). The performance of RMSEA in models with small degrees of freedom. *Sociological Methods & Research*, *44*(3), 486-507. https://doi.org/10.1177/0049124114543236

Kieslich, K., Lünich, M., & Marcinkowski, F. (2021). The Threats of Artificial Intelligence Scale (TAI): development, measurement and test over three application domains. *International Journal of Social Robotics, 13*(7), 1563-1577. https://doi.org/10.1007/s12369-020-00734-w

Kroenke, K., Spitzer, R. L., & Williams, J. B. W. (2003). The Patient Health Questionnaire-2: Validity of a two-item depression screener. *Medical Care, 41*(11), 1284–1292. https://doi.org/10.1097/01.MLR.0000093487.78664.3C

Kucherenko, O.O. (2025). A scoping review of psychometric scales assessing attitudes. toward artificial intelligence. *Technologies of Intellect Development, 9*(2), 1-27. https://doi.org/10.31108/3.2025.9.2.18

Kuru, O., & Pasek, J. (2016). Improving social media measurement in surveys: Avoiding acquiescence bias in Facebook research. *Computers in Human Behavior, 57*, 82-91. https://doi.org/10.1016/j.chb.2015.12.008

Lee, S. Y., Kim, H. M., & Ahn, H.-N. (2025). Validation of the Korean version of the Attitudes Toward Artificial Intelligence Scale (ATTARI-12). *Korean Journal of Culture and Social Issue, 31*(2), 269-292. https://doi.org/10.20406/kjcs.2025.5.31.2.269

Li, C. H. (2016). Confirmatory factor analysis with ordinal data: Comparing robust maximum likelihood and diagonally weighted least squares. *Behavior Research Methods*, *48*(3), 936-949. https://doi.org/10.3758/s13428-015-0619-7

Li, P., Chong, Y. Y., Man, W., & Li, H. (2026). Measuring attitudes toward artificial intelligence: validation of the Chinese ATTARI-12 and invariance across genders and academic disciplines. *Frontiers in Psychology*, *17*, Article 1836060. https://doi.org/10.3389/fpsyg.2026.1836060

Maier, R., Ghiocanu S. C., Costin, A., & Simion, I. (2025). Students' attitude towards artificial intelligence and its predictive factors. *Broad Research in Artificial Intelligence and Neuroscience, 16*(1). https://doi.org/10.70594/brain/16.S1/22

Malecki, W. P., Messingschlager, T. V., & Appel, M. (2025). The impact of exposure to generative AI art on aesthetic appreciation, perceptions of AI mind, and evaluations of AI and of art careers. *New Media & Society*, *27*(9), 5410-5432. https://doi.org/10.1177/14614448251344590

Maslej, N., Fattorini, L., Perrault, R., Gil, Y., Parli, V., …, & Oak, S. (2025). *The AI index 2025 annual report*. AI Index Steering Committee, Institute for Human-Centered AI, Stanford University. https://hai.stanford.edu/ai-index/2025-ai-index-report

Messick, S. (1995). Validity of psychological assessment: Validation of inferences from persons' responses and performances as scientific inquiry into score meaning. *American Psychologist, 50*(9), 741–749. https://doi.org/10.1037/0003-066X.50.9.741

Montag, C., Nadaf, Y., & Kamble, S. V. (2026). Investigating global AI attitudes, experience with AI and trust in ChatGPT in a sample from India: A study attempt with preliminary findings. *Open Psychology*, *8*(1), Article 20250012. https://doi.org/10.1515/psych-2025-0012

Nye, C. D., Bradburn, J., Olenick, J., Bialko, C., & Drasgow, F. (2019). How big are my effects? Examining the magnitude of effect sizes in studies of measurement equivalence. *Organizational Research Methods, 22*, 678–709. https://doi.org/10.1177/1094428118761122

R Core Team (2025). *R: A Language and Environment for Statistical Computing* (Version 4.5.1) [Computer software]. Austria: R Foundation for Statistical Computing. https://www.R-project.org

Rolstad, S., Adler, J., & Rydén, A. (2011). Response burden and questionnaire length: is shorter better? A review and meta-analysis. *Value in Health*, *14*(8), 1101-1108. https://doi.org/10.1016/j.jval.2011.06.003

Rosenberg, M. J., & Hovland, C. I. (1960). *Attitude organization and change: An analysis of consistency among attitude components.* Yale University Press.

Rosseel, Y. (2012). lavaan: An R package for structural equation modeling. *Journal of Statistical Software, 48*(2), 1-36. https://doi.org/10.18637/jss.v048.i02

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

Schepman, A., & Rodway, P. (2025). The measurement of attitudes towards artificial intelligence: An overview and recommendations. In C. Montag and R. Ali (Eds.), *The*

*impact of artificial intelligence on societies: Understanding attitude formation towards AI* (pp. 9-24). Springer.

Schepman, A., & Rodway, P. (2026). Validation of the Short General Attitudes Towards Artificial Intelligence Scale: The Short GAAIS-10. *International Journal of Human–Computer Interaction*. Advance online publication. https://doi.org/10.1080/10447318.2025.2610446

Schermelleh-Engel, K., Moosbrugger, H., & Müller, H. (2003). Evaluating the fit of structural equation models: tests of significance and descriptive goodness-of-fit measures. *Methods of Psychological Research, 8*(2), 23–74. https://doi.org/10.23668/psycharchives.12784

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, G. T., McCarthy, D. M., & Anderson, K. G. (2000). On the sins of short-form development. *Psychological Assessment*, *12*, 102-111. https://doi.org/10.1037//1040-3590.12.1.102

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

Stein, J.-P. & Ohler, P. (2017). Venturing into the uncanny valley of mind – The influence of mind attribution on the acceptance of human-like characters in a virtual reality setting. *Cognition, 160,* 43–50. https://doi.org/10.1016/j.cognition.2016.12.010

Suh, W., & Ahn, S. (2022). Development and validation of a scale measuring student attitudes toward artificial intelligence. *Sage Open*, *12*(2). https://doi.org/10.1177/21582440221100463

Svetina, D., Rutkowski, L., & Rutkowski, D. (2020). Multiple-group invariance with categorical outcomes using updated guidelines: An illustration using *Mplus* and the *lavaan/semTools* packages. *Structural Equation Modeling*, *27*(1), 111-130. https://doi.org/10.1080/10705511.2019.1602776

Steyer, R., Schmitt, M., & Eid, M. (1999). Latent state-trait theory and research in personality and individual differences. *European Journal of Personality, 13*, 389-408. https://doi.org/10.1002/(SICI)1099-0984(199909/10)13:5<389::AID-PER361>3.0.CO;2-A

Vorobeva, D., El Fassi, Y., Costa Pinto, D., Hildebrand, D., Herter, M. M., & Mattila, A. S. (2022). Thinking skills don't protect service workers from replacement by artificial intelligence. *Journal of Service Research, 25*(4), 601–613. https://doi.org/10.1177/10946705221104312

Walker, J. (2019). Non-inferiority statistics and equivalence studies. *BJA Education, 19*(8), 267-271. https://doi.org/10.1016/j.bjae.2019.03.004

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

Weijters, B., & Baumgartner, H. (2022). On the use of balanced item parceling to counter acquiescence bias in structural equation models. *Organizational Research Methods*, *25*(1), 170-180. https://doi.org/10.1177/1094428121991909

Weinerová, J., Szücs, D., & Ioannidis, J. (2022). Published correlational effect sizes in social and developmental psychology. *Royal Society Open Science*, *9*(12), Article 220311. https://doi.org/10.1098/rsos.220311

Weinmann, H., Messingschlager, T.V., & Appel, M. (2026). Gender bias in text-to-image generative artificial intelligence: neglect and stereotypical presentations across three popular platforms. *New Media & Society.* Advanced online publication. https://doi.org/10.1177/14614448261435197

Wu, H., & Estabrook, R. (2016). Identification of confirmatory factor analysis models of different levels of invariance for ordered categorical outcomes. *Psychometrika*, *81*(4), 1014-1045. https://doi.org/10.1007/s11336-016-9506-0

Zanna, M. P., & Rempel, J. K. (2008). Attitudes: A new look at an old concept. In R. H. Fazio & R. E. Petty (Eds.), *Attitudes: Their structure, function, and consequences* (pp. 7–15). Psychology Press.

Ziegler, M., Poropat, A., & Mell, J. (2014). Does the length of a questionnaire matter?. *Journal of Individual Differences*, *35*(4), 250-261. https://doi.org/10.1027/1614-0001/a000147

**Table 1**

*Item Characteristics of the ATTARI-12 items as observed in previous work (Stein et al., 2024, Study 1) and Short Form Inclusion*

| | English Version Item Wording | Facet | P/N | Pop. | General factor loading | | ATTARI-6 | ATTARI-3 |
|---|---|---|---|---|---|---|---|---|
| | | | | | Unstandardized | Standardized | | |
| 1 | AI will make this world a better place. | C | P | .66 | 0.79 (0.04) | .84 | x | x |
| 2 | I have strong negative emotions about AI. | A | N | .73 | 0.66 (0.05) | .59 | x | |
| 3 | I want to use technologies that rely on AI. | B | P | .61 | 0.81 (0.04) | .77 | x | x |
| 4 | AI has more disadvantages than advantages. | C | N | .65 | 0.66 (0.05) | .57 | | |
| 5 | I look forward to future AI developments. | A | P | .70 | 0.91 (0.04) | .87 | | |
| 6 | AI offers solutions to many world problems. | C | P | .70 | 0.64 (0.05) | .65 | | |
| 7 | I prefer technologies that do not feature AI. | B | N | .62 | 0.77 (0.05) | .64 | x | |
| 8 | I am afraid of AI. | A | N | .75 | 0.55 (0.05) | .48 | | |
| 9 | I would rather choose a technology with AI than one without it. | B | P | .53 | 0.71 (0.05) | .65 | | |
| 10 | AI creates problems rather than solving them. | C | N | .67 | 0.67 (0.05) | .59 | x | |
| 11 | When I think about AI, I have mostly positive feelings. | A | P | .65 | 0.95 (0.04) | .88 | x | x |
| 12 | I would rather avoid technologies that are based on AI. | B | N | .69 | 0.84 (0.05) | .69 | | |

*Note*. *A* = Affective Facet, *B* = Behavioral Facet, *C* = Cognitive Facet, *N* = Negatively Worded, *P* = Positively Worded. Pop. = Standardized item popularity. For the model specification, please refer to Stein et al. (2024). Standard errors of the unstandardized factor loadings are in parentheses.

**Table 2**

*Means, Standard Deviations, and Factor Loadings for ATTARI-6 and ATTARI-3 Items*

| | | | | | | ATTARI-6 | | | | ATTARI-3 | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | Item | Facet | *M* | *SD* | $r_{it}$ | $\lambda_g$ | $\lambda_{g(m)}$ | $\lambda_m$ | ECV | $r_{it}$ | $\lambda_g$ |
| *English version* | | | | | | | | | | | |
| 1. | Artificial intelligence will make this world a better place. | Cognitive | 3.49 | 1.04 | .80 | .89 | .90 | | 1.00 | .80 | .90 |
| 2. | I have strong negative emotions about artificial intelligence. * | Affective | 3.81 | 1.12 | .71 | .79 | .73 | .39 | 0.78 | | |
| 3. | I want to use technologies that rely on artificial intelligence. | Behavioral | 3.32 | 1.09 | .71 | .81 | .82 | | 1.00 | .74 | .84 |
| 7. | I prefer technologies that do not feature artificial intelligence. * | Behavioral | 3.32 | 1.19 | .72 | .79 | .72 | .43 | 0.74 | | |
| 10. | Artificial intelligence creates problems rather than solving them. * | Cognitive | 3.69 | 1.07 | .73 | .81 | .76 | .34 | 0.83 | | |
| 11. | When I think about artificial intelligence, I have mostly positive feelings. | Affective | 3.50 | 1.13 | .83 | .92 | .94 | | 1.00 | .81 | .93 |
| *German version* | | | | | | | | | | | |
| 1. | Künstliche Intelligenz wird die Welt verbessern. | Cognitive | 3.33 | 0.97 | .68 | .78 | .79 | | 1.00 | .69 | .83 |
| 2. | Ich habe starke negative Emotionen gegenüber künstlicher Intelligenz. * | Affective | 3.84 | 1.11 | .70 | .80 | .76 | .33 | 0.84 | | |
| 3. | Ich möchte Technologien nutzen, die auf künstlicher Intelligenz basieren. | Behavioral | 3.55 | 1.02 | .72 | .79 | .80 | | 1.00 | .68 | .79 |
| 7. | Ich bevorzuge Technologien, die keine künstliche Intelligenz beinhalten. * | Behavioral | 3.24 | 1.14 | .64 | .71 | .69 | .17 | 0.95 | | |
| 10. | Künstliche Intelligenz verursacht eher Probleme, anstatt sie zu lösen. * | Cognitive | 3.52 | 1.00 | .70 | .79 | .75 | .28 | 0.88 | | |
| 11. | Wenn ich an künstliche Intelligenz denke, habe ich hauptsächlich positive Gefühle. | Affective | 3.10 | 1.03 | .77 | .87 | .88 | | 1.00 | .72 | .86 |

*Note*. $N$ = 1,023 / 994 for English and German version. The item numbers correspond to the ATTARI-12 (Stein et al., 2024). Negatively worded items are marked with an asterisk. $r_{it}$ = Corrected item-total correlation; $\lambda_g$ = Standardized loading on general factor in general factor model; $\lambda_{g(m)}$ = Standardized loading on general factor in bifactor factor model; $\lambda_m$ = Standardized loading on method factor in bifactor factor model. ECV = Explained common variance by general factor.

**Table 3**

*Means, Standard Deviations, Reliabilities, and Correlations of the ATTARI Scales*

| Scale | *M* | *SD* | Skewness | Kurtosis | α | ω | Correlations | |
|---|---|---|---|---|---|---|---|---|
| | | | | | | | ATTARI-6 | ATTARI-3 |
| *English version* | | | | | | | | |
| ATTARI-12 | 3.56 | 0.88 | -0.52 | -0.20 | .94 | .96 | .98* | .95* |
| ATTARI-6 | 3.52 | 0.92 | -0.54 | -0.12 | .91 | .92 | | .97* |
| ATTARI-3 | 3.53 | 0.94 | -0.58 | -0.06 | .89 | .89 | | |
| *German version* | | | | | | | | |
| ATTARI-12 | 3.52 | 0.77 | -0.64 | 0.08 | .93 | .95 | .98* | .95* |
| ATTARI-6 | 3.43 | 0.83 | -0.65 | 0.15 | .89 | .89 | | .97* |
| ATTARI-3 | 3.45 | 0.86 | -0.69 | 0.19 | .83 | .83 | | |

*Note*. $N$ = 1,023 / 994 for English and German version. α / ω = Coefficient alpha / omega reliability.
* $p < .05$.

**Table 4**

*Model Fit Statistics for Measurement Models of the ATTARI-6*

| Model | $\chi^2$ (*df*) | CFI | TLI | SRMR | $\Delta\chi^2$ (*df*) | *p*-value |
|---|---|---|---|---|---|---|
| *English version* | | | | | | |
| 1. Single factor model | 373.63 (9) | .98 | .95 | .04 | | |
| 2. Bifactor facet model | 328.57 (7) | .98 | .95 | .04 | 0.59 (2) | .746 |
| 3. Bifactor method model | 54.49 (6) | 1.00 | .99 | .01 | 236.47 (3) | < .001 |
| *German version* | | | | | | |
| 1. Single factor model | 124.72 (9) | .99 | .98 | .03 | | |
| 2. Bifactor facet model | 128.99 (7) | .99 | .97 | .03 | 2.23 (2) | .328 |
| 3. Bifactor method model | 101.51 (6) | .99 | .97 | .03 | 31.35(3) | < .001 |

*Note*. $N$ = 1,023 / 994 for English and German version. CFI = Comparative fit index, TLI = Tucker-Lewis index, SRMR = Standardized root mean residual. Visual representations of the models and factor loadings are given in Figure S2 of the supplementary material.

**Table 5**

*Means, Standard Deviations, and Factor Loadings for ATTARI-6 and ATTARI-3 Items in Study 2*

| | | *ATTARI-6* | | | | | | ATTARI-3 | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Item | Facet | *M* | *SD* | $r_{it}$ | $\lambda_g$ | $\lambda_{g(m)}$ | $\lambda_m$ | *M* | *SD* | $r_{it}$ | $\lambda_g$ |
| 1. | Cognitive | 2.99 | 1.21 | .84 | .91 | .92 | | 3.14 | 1.19 | .85 | .93 |
| 2. | Affective | 3.39 | 1.36 | .84 | .90 | .87 | .34 | | | | |
| 3. | Behavioral | 2.91 | 1.28 | .82 | .88 | .89 | | 3.17 | 1.24 | .84 | .90 |
| 7. | Behavioral | 2.97 | 1.33 | .84 | .90 | .86 | .33 | | | | |
| 10. | Cognitive | 3.27 | 1.21 | .80 | .85 | .81 | .29 | | | | |
| 11. | Affective | 3.03 | 1.28 | .88 | .94 | .96 | | 3.18 | 1.27 | .88 | .97 |

*Note*. $N$ = 396 / 374 for ATTARI-6 and ATTARI-3. The item numbers correspond to the ATTARI-12 (Stein et al., 2024). $r_{it}$ = Corrected item-total correlation; $\lambda_g$ = Standardized loading on general factor in general factor model; $\lambda_{g(m)}$ = Standardized loading on general factor in bifactor factor model; $\lambda_m$ = Standardized loading on method factor in bifactor factor model.

**Table 6**

*Correlations Between the ATTARI Versions with the Validity Indicators*

| | *Zero-order Correlations with 95% Confidence Interval* | | | *Difference in Correlations with 95% Confidence Interval* | |
|---|---|---|---|---|---|
| Variable | ATTARI-12 | ATTARI-6 | ATTARI3 | ATTARI-6 – ATTARI-12 | ATTARI3 – ATTARI-12 |
| Agreeableness | .25 [.15, .35] | .19 [.08, .29] | .29 [.19, .38] | -.06* [-.20, .09] | .04* [-.11, .18] |
| Conspiracy Mentality | –.10 [-.21, .01] | –.10 [-.21, .01] | –.06 [-.17, .04] | .00* [-.15, .15] | .04* [-.12, .19] |
| Career Interest | .83 [.78, .86] | .86 [.83, .88] | .84 [.80, .87] | .03* [-.02, .08] | .01* [-.04, .06] |
| AI Anxiety (Job Replacement) | –.58 [-.64, -.49] | –.65 [-.70, -.59] | –.50 [-.58, -.38] | -.07* [-.17, .02] | .08* [-.04, .21] |
| Eeriness Response | –.58 [-.65, -.50] | –.48 [-.56, -.39] | –.39 [-.48, -.28] | .10* [-.01, .22] | .20* [.08, .32] |
| Age | .13 [.03, .22] | .17 [.07, .26] | .17 [.08, .26] | .04* [-.10, .17] | .04* [-.09, .18] |

*Note*. $N$ = 384 / 396 / 374 for ATTARI-12, ATTARI-6, and ATTARI-3. Bias-corrected and accelerated confidence intervals based on 5,000 bootstrap samples.

* $p < .05$ for non-inferiority test

# APPENDIX

## Table A1

ATTARI-6 and ATTARI-3: Instruction and Item Wordings

*In the following, we are interested in your attitudes towards artificial intelligence (AI). AI can execute tasks that typically require human intelligence. It enables machines to sense, act, learn, and adapt in an autonomous, human-like way. AI may be part of a computer or online platform—but it can also be encountered in various other hardware devices such as robots.*

| ATTARI-6 | ATTARI-3 | Item wordings | |
|---|---|---|---|
| | | ***English version*** | ***German version*** |
| 1. / C / p | 1. / C / p | **Artificial intelligence will make this world a better place.** | **Künstliche Intelligenz wird die Welt verbessern.** |
| 2. / A / n | | I have strong negative emotions about artificial intelligence. | Ich habe starke negative Emotionen gegenüber künstlicher Intelligenz. |
| 3. / B / p | 2. / B / p | **I want to use technologies that rely on artificial intelligence.** | **Ich möchte Technologien nutzen, die auf künstlicher Intelligenz basieren.** |
| 4. / B / n | | I prefer technologies that do not feature artificial intelligence. | Ich bevorzuge Technologien, die keine künstliche Intelligenz beinhalten. |
| 5. / C / n | | Artificial intelligence creates problems rather than solving them. | Künstliche Intelligenz verursacht eher Probleme, anstatt sie zu lösen. |
| 6. / A / p | 3. / A / p | **When I think about artificial intelligence, I have mostly positive feelings.** | **Wenn ich an künstliche Intelligenz denke, habe ich hauptsächlich positive Gefühle.** |

*Note*. New item numbering. C = Cognitive, A = Affective, B = Behavioral, *p* = positively worded, *n* = negatively worded. An item wording in bold designates that the item is part of the ATTARI-3 scale. Prior to building the ATTARI-6 mean score, the item scores of the negatively worded items need to be reversed.