Developing and Validating the Chinese Version of the Attitudes Toward Large Language Models Scale (AT-LLM Chinese)

Sameha AlShakhsi

salshakhsi@hbku.edu.qa

Hamad bin Khalifa University

Ala Yankouskaya

Bournemouth University

Haibo Yang

Tianjin Normal University

Xiaokun Wang

University of Science and Technology Beijing

Jiaojiao Chen

Tianjin Normal University

Tina Yunsi MA

Education University of Hong Kong

Guandong Xu

Education University of Hong Kong

Christian Montag

University of Macau

Raian Ali

Hamad bin Khalifa University

Research Article

Keywords: Attitude, Large Language Models, Scale, Generative AI, Human-AI Collaboration

Posted Date: January 14th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-8457517/v1

Additional Declarations: No competing interests reported.

Abstract

Large Language Models (LLMs) have become integral to education, business, and public life, yet cross-culturally validated tools for assessing public attitudes toward them remain scarce. This study adapted and validated two established five-item instruments, the Attitudes Toward General LLMs (AT-GLLM) and Attitudes Toward Primary LLMs (AT-PLLM) scales, for use in the Chinese context. Each scale includes two items measuring acceptance and three items measuring fear. A sample of 576 Chinese LLM users completed the Chinese versions of both scales alongside the Attitudes Toward Artificial Intelligence (ATAI) measure and a self-efficacy scale. Confirmatory factor analyses supported the expected two-factor structure, acceptance and fear, for both scales, with acceptable model fit indices. Reliability coefficients ranged from α = .54 to .74, with the lower value corresponding to the two-item acceptance subscale, as expected given its brevity. Measurement invariance testing across low- and high-frequency LLM users confirmed configural, metric, scalar, and strict invariance, indicating that the constructs operate equivalently across experience levels. External validation showed that ATAI-acceptance strongly predicted LLM acceptance, whereas ATAI-fear predicted LLM-related fear, supporting convergent validity. Self-efficacy did not significantly predict LLM attitudes once general AI attitudes were accounted for. These findings confirm the psychometric soundness and cross-cultural applicability of the AT-GLLM and AT-PLLM scales in China. By providing validated instruments for one of the world's largest and most active AI ecosystems, this study advances global understanding of LLM attitudes and offers tools for guiding responsible, trust-oriented LLM design and policy development.

1. Introduction

The rise of large language models (LLMs) such as ChatGPT and DeepSeek has ushered in a new era of generative artificial intelligence (AI). LLMs have transformed the way humans interact with AI, from systems that simply follow programmed instructions to tools capable of engaging in dynamic, human-like conversations. Their rapid advancement has led to deep integration across multiple sectors, including education, healthcare, creative industries, and public administration. In China, this development is particularly significant, not only in terms of adoption but also innovation, as domestic LLMs have proliferated rapidly, supported by national policies promoting technological self-reliance and AI sovereignty (Chang et al., 2025). China's dual role as both a major user and a competitive developer of LLMs has created a unique socio-technical landscape, making public attitudes toward these systems an important area of investigation.

As with artificial intelligence (AI) more broadly, LLMs, as a subset of AI, have elicited both enthusiasm and concern. They are widely regarded as innovative technologies that introduce novel capabilities beyond existing digital tools, yet they also raise important ethical, social, and practical questions. This duality is evident in the Chinese context as well. For instance, a sentiment analysis of Chinese social media posts revealed generally positive public attitudes toward ChatGPT, while simultaneously revealing persistent concerns regarding ethical risks, misinformation, job displacement, and human–computer relationship (Lian et al., 2024). Similarly, Chinese oncologists reported both optimism about AI-driven

chatbots improving accessibility and concerns over misinformation, liability, lack of personalization, and privacy (Zeng et al., 2025). A cross-sectional study among Chinese medical students found that most participants exhibited neutral or moderately positive attitudes toward LLMs and perceived them as beneficial for medical (Pan & Ni, 2024). Therefore, understanding public attitudes toward LLMs is critical, because attitudes shape technology adoption, trust, and regulation, all of which impact how these systems integrate into society. In the Chinese context, the stakes are high: user acceptance influences not only consumer uptake but also the legitimacy of state-led AI initiatives and the sociocultural alignment of technology.

To develop responsible and context-sensitive LLM-based technologies and to support institutions in navigating the complexities of integrating LLM-driven tools across diverse domains, understanding the factors that influence LLM adoption is essential. Attitudes toward LLMs, including trust, perceived usefulness and ethical acceptance, represent a key determinant of adoption behavior. Scholars have emphasized that realizing the potential of LLMs requires frameworks grounded in trust, transparency, fairness, and responsible AI governance (Sarker, 2024). In educational settings, a national study in higher education institutions found that trust significantly influenced students' perceptions of ChatGPT's usefulness and ease of use, thereby shaping their intention to adopt it for academic purposes (Shahzad et al., 2024). Similarly, in medical contexts, a study of Chinese doctors identified perceived usefulness, technical support, and social influence as key determinants of LLM adoption, while perceived risk, including ethical concerns and misinformation, negatively influenced usage behavior (Qu et al., 2024). Furthermore, a study developing medical chatbots highlight that ethical and privacy concerns directly affect users' satisfaction and loyalty toward continued use (Niu & Mvondo, 2024). Despite growing empirical research on attitudes and adoption factors, there remains limited work on systematically measuring these attitudes. Developing reliable and valid scales to assess public and professional attitudes toward LLMs is therefore critical for advancing empirical understanding and guiding ethical, user-centered AI integration.

While numerous studies have developed instruments to assess attitudes toward artificial intelligence (AI) (Schepman & Rodway, 2020; Stein et al., 2024), research specifically measuring attitudes toward LLMs remains limited. For instance, a recent study used Technology Acceptance Model-Based Scale (TAME-ChatGPT) to examine nursing students' attitudes toward LLMs in relation to their metacognitive abilities (Yeh & Siah, 2025). Survey-based research among internal medicine residents has also explored concerns regarding LLM use using structured survey instruments that included items on hallucinations, ethical implications, and misinformation (Fried et al., 2024). Studies have also used qualitative and mixed-method investigations to explore user perceptions of LLMs in practice. A study used mixed-methods sentiment and thematic analysis found that users have a highly positive attitude toward DeepSeek, valuing its free analytical power but noting ethical and usability including censorship, data privacy, and interface issues (Albuhairy & Algaraady, 2025). Another study used qualitative analysis reported a positive attitude toward DeepSeek's reliability and usefulness for language learning, though ethical and privacy concerns persisted (Habeb Al-Obaydi & Pikhart, 2025). Other research has modified existing attitudes or applied frameworks to assess attitude toward LLMs. One study employed a

modified version of the General Attitudes Toward Artificial Intelligence Scale (GAAIS) to explore perceptions among students and faculty, finding that while students expressed enthusiasm and a willingness to engage with LLMs, faculty participants raised greater concerns about ethics, transparency, reliability, and the potential erosion of human interaction in education (Rajik, 2024). Zhao et al. (2024) applied the Technology Acceptance Model (TAM) and Innovation Diffusion Theory (IDT) to assess educators' attitudes toward LLMs. Similarly, a multinational survey used the TAM framework to examine user intent to adopt DeepSeek for healthcare purposes, revealing that over half of participants were likely to switch to DeepSeek over other LLMs (Choudhury et al., 2025). Finally, recent work showed that more positive global attitudes towards AI translate more to trust in ChatGPT in German samples (Montag & Ali, 2023) and more trust in ChatGPT/Ernie Bot in a Chinese sample, but associations were far away from perfect (Montag et al., 2025). This suggests that attitudes may differ between general perceptions of LLMs/AI and attitudes toward specific models.

Although existing approaches capture general attitudes, ranging from optimism and trust to concerns about bias and ethical implications, they lack systematic and psychometrically validated measurement. Moreover, further evidence is needed to determine whether instruments originally developed to assess attitudes toward AI can be reliably adapted for LLMs, which represent a distinct subclass of AI technologies. This limitation stems from the fundamental difference between traditional AI systems, which often operate in the background of user experience, and LLMs, which are interactive, conversational, and anthropomorphized.

To address this gap, Liebherr et al. (Liebherr et al., 2025) and Barajeeh et al (2025a) validated two dedicated instruments, the Attitudes Toward General LLMs (AT-GLLM) and Attitudes Toward Primary LLM (AT-PLLM) scales, in English and Arabic, respectively. The English version was validated with a UK sample, while the Arabic version was tested with an Arab sample. These scales adapt the five-item ATAI scale (Sindermann et al., 2021) to the LLM context, measuring both acceptance (two items reflecting trust and perceived benefit) and fear (three items reflecting ethical concerns, job loss, and existential risk). The AT-GLLM assesses broad societal attitudes, while the AT-PLLM captures more personal, experience-based evaluations toward one's most frequently used model (e.g., ChatGPT or Deepseek). Both scales showed solid psychometric validity, reliability in UK and Arab samples. Given China's sociotechnical environment, linguistic diversity, and government-led AI initiatives, it is crucial to test and adapt these instruments for use in the Chinese context. The present work addresses this gap by examining the applicability and cultural relevance of LLM attitude measures in China, contributing to a more globally inclusive understanding of how people perceive and engage with generative AI technologies.

The present study aims to extend prior work by adapting and validating psychometric instruments designed to assess attitudes toward LLMs within the Chinese context. Building on the Attitudes Toward Artificial Intelligence (ATAI) scale, we modify item wording to ensure relevance to LLM-specific experiences, for example, replacing general references to "AI" with "LLMs" or "my primary LLM.", while maintaining the underlying structure and conceptual continuity of the original instrument. Two versions

of the adapted scale are tested: the Attitudes Toward General LLMs (AT-GLLM), and the Attitudes Toward Primary LLM (AT-PLLM). Using a large Chinese sample, we conduct exploratory and confirmatory factor analyses to examine the factorial structure and measurement invariance of both instruments. In doing so, this study provides a systematic validation of LLM attitude measures in China, contributing to a contextual understanding of how users in one of the world's leading AI ecosystems perceive and engage with generative AI technologies. Furthermore, we expect that the AT-GLLM will demonstrate a strong association with the original ATAI, as both instruments assess generalized attitudes toward AI technologies.

2. Method

2.1. Study Design

The items used in this study were adapted from the English version of the AT-GLLM and AT-PLLM scales proposed by Liebherr et al. (2025) for use with the Chinese sample. These scales were created as part of a broader research project investigating various factors related to LLM usage, including attitudes toward LLMs, levels of dependency on LLMs, and associations with personal and usage-related factors. In the present analysis, only the sections of the questionnaire relevant to the development of the Chinese versions of the AT-GLLM and AT-PLLM scales, as well as the variables employed for their external validation, were utilized.

At the beginning of the survey, participants were presented with a brief introduction of LLMs, including their scope and functionality. It clarifies that LLMs go beyond conversational systems such as ChatGPT and include a broader range of functions and applications. This introduction ensures that all participants have a similar baseline knowledge and ensure consistency in their responses regarding their familiarity and use of LLMs. The introduction presented to participants was as follows:

"Large Language Models (LLMs) are advanced artificial intelligence systems designed to process and generate human-like text by leveraging deep learning techniques and extensive datasets. A subset of these models, known as conversational LLMs, are specifically designed to engage in natural, interactive dialogues with users. These models are trained on vast amounts of text data, enabling them to generate meaningful, context-aware replies during interactions.

Conversational LLMs are versatile and can assist with tasks such as answering questions, generating ideas, or engaging in casual dialogue. Examples include ChatGPT by OpenAI, Gemini by Google, DeepSeek by DeepSeek AI, Claude by Anthropic, Bing Chat by Microsoft, and Ernie Bot by Baidu."

The following section gathered participant's demographic information including age, gender, employment status, and education levels. To measure participants' engagement with LLMs, the survey included an item asking respondents to rate their frequency of LLM use on a scale from 0 to 10. The subsequent section assessed participants' attitudes toward LLMs, both in general and the primary one, attitude toward artificial intelligence, using the Attitudes Toward Artificial Intelligence (ATAI)

questionnaire. To reduce potential response bias and prevent habituation effects, the three questionnaires were not presented consecutively, ensuring greater independence in participants' responses. Following this, participants completed a single-item measure of self-efficacy, also rated on a 0–10 scale.

2.2. Sample

Participants were recruited via Credemo and WJX, two online data collection platforms. The sample comprised individuals aged between 18 and 60 years who were residing in China. Data were screened for quality, including the removal of incomplete responses and participants who failed attention checks. The final sample consisted of 576 participants. participants were required to have prior knowledge of LLMs and to have used at least one LLM-based website or application. Participants were also asked to identify their most frequently used LLM (referred to as their "Primary LLM") and to report how often they used it. To ensure that participants were actual users of LLMs, they rated their usage frequency on a scale from 0 to 10, those who selected 0 were excluded.

The adequacy of the sample size was evaluated against established criteria for questionnaire validation using confirmatory factor analysis (CFA). Each scale comprised five items specified in a two-factor model. Methodological guidance indicates that CFA models with fewer than six indicators and simple factor structures are adequately powered for parameter estimation and model fit evaluation when sample size exceeds 200-300 observations (MacCallum et al., 1999). Simulation studies further show that samples above 500 yield low bias, high solution propriety, and stable fit indices for small CFA and SEM models with moderate-to-high loadings (Wolf et al., 2013). The achieved sample size therefore exceeds commonly cited thresholds for reliable estimation in models of this size and complexity. In addition, the study examined measurement invariance across frequency-of-use groups defined by participants' self-reported frequency of use of their primary LLM (low-frequency vs high-frequency users). Guidance on multi-group CFA indicates that invariance testing across two groups remains identifiable with unequal group sizes when the measurement model is simple and each group contains at least 100 observations (Putnick & Bornstein, 2016). The group sizes in the present study satisfy these conditions, supporting the estimation of configural, metric, scalar, and strict invariance across usage-frequency groups.

Ethical approval for the study was obtained from Bournemouth University, UK (N62239, 03.03.2025). Participants were provided with an information sheet detailing the study's aims, procedures, data handling, and withdrawal rights. Informed consent was obtained electronically prior to survey commencement.

2.3. Measures

Attitude towards Artificial Intelligence (ATAI)

Participants' attitude toward AI was measured using a five-item scale developed by Sindermann et al. (2021). Each item of the scale was rated on an 11-point Likert scale from 0 (Strongly Disagree) to 10 (Strongly Agree). The ATAI includes two negatively correlated dimensions: acceptance measured through trust and perceived benefits (two items), and fear measured through fear of AI, perceived risk to humanity, and concerns about job loss (three items).

Attitude Toward General Large Language Models (AT-GLLM)

This instrument evaluates participants' attitudes toward general LLMs. To guide participants, the questionnaire was preceded by a brief instruction clarifying its purpose. Participants were instructed: "*Think of LLMs in general and the applications built on them. To what extent does the following statement apply to you?*" Items were rated on a scale from 0 (strongly disagree) to 10 (strongly agree). The scale retained the original structure and content of the ATAI measure, but referents were adjusted to reflect the broader scope of this version through minor rewording and the introductory instruction. For example, the ATAI item "I trust artificial intelligence" was modified to "I trust them" to align with the general framing of LLMs. The scale comprises two dimensions: acceptance (two items, e.g., "I trust them") and fear (three items, e.g., "I fear them"). Table 1 lists all items in both English and Chinese.

Attitude Toward Primary Large Language Models (AT-PLLM)

This scale includes five items assessing participants' attitudes toward their primary or most frequently used LLM. Before responding, participants were presented with the following instruction: "Think of your primary LLM, meaning the LLM you use the most. To what extent does the following statement apply to you?". Each item was rated on the same 0–10 Likert-type used for the ATAI and AT-GLLM. AT-PLMM captures two key aspects of attitude: acceptance (two items, e.g., "I trust it") and fear (three items, e.g., "I fear it").

Table 1. Items of the AT-GLLM and AT-PLLM Scales: English and Chinese Versions

		English	Chinese
AT-GLLM	Attitude toward General LLMs	I fear them	我畏惧它
		I trust them	我信任它
		They will destroy humankind	它将毁灭人类
		They will benefit humankind	它将造福人类
		They will cause many job losses	它会导致大量的失业
AT-PLLM	Attitude toward Primary LLM	I fear it	我畏惧它
		I trust it	我信任它
		It will destroy humankind	它将毁灭人类
		It will benefit humankind	它将造福人类
		it will cause many job losses	它会导致大量的失业

Self-Efficacy

Self-efficacy was measured using the Single-Item General Self-Efficacy Scale (GSE-SI) (Di et al., 2023) as external criterion for validation. Participants rated the item, “I believe I can succeed at most of any endeavour to which I set my mind?” using an 11-point Likert scale (0= Not true of me to 10 = Very True of Me). Higher scores indicated greater levels of self-efficacy.

2.4. Data Analysis

The data were analysed using a multi-step validation process to examine the psychometric properties of the adapted AT-GLLM and AT-PLLM scales.

Monotonicity check. The assumption of monotonicity between latent traits and item responses was examined using nonparametric isotonic regression, which offers a robust and flexible approach that does not rely on distributional assumptions. For each item, a restscore was calculated by summing responses to all other items within the same scale, serving as a predictor of the target item’s response. The fitted regression function was inspected to verify whether the expected item value increased monotonically with the restscore. Results indicated that all items in both the AT-GLLM and AT-PLLM scales satisfied the monotonicity assumption. The fitted values ranged from 0 to 10, confirming a consistent non-decreasing relationship between the latent trait and item responses. In the AT-GLLM, the trust item and benefit items displayed more moderate increases, whereas the fear and destruction concerns items showed the steepest increase. In AT-PLLM, a similar pattern was observed, with fear and

destruction concern items showing the steepest slopes and the benefit item exhibiting a more moderate increase.

Confirmatory Factor Analysis. Confirmatory factor analysis (CFA) was conducted to verify the factor structure of the AT-GLLM and AT-PLLM scales based on prior work (Barajeeh et al., 2025a; Liebherr et al., 2025). The analyses were performed in R using the *lavaan* package (Rosseel, 2012), applying maximum likelihood estimation with robust standard errors (MLR). Model fit was assessed using several indices: the Chi-square statistic (χ^2), Comparative Fit Index (CFI), Tucker–Lewis Index (TLI), Root Mean Square Error of Approximation (RMSEA), and Standardized Root Mean Square Residual (SRMR). A good model fit was indicated by CFI and TLI values > .90 and RMSEA and SRMR values < .08. The results supported the hypothesized two-factor structure for both scales, representing acceptance and fear dimensions of attitudes toward LLMs.

Reliability analysis. Internal consistency and construct reliability were assessed using several complementary indices to capture both classical and latent-variable-based reliability. Analyses were conducted separately for the fear and acceptance subscales of each questionnaire (ATAI, attitudes toward general LLMs, and attitudes toward primary LLMs).

Cronbach's alpha (α) was used as a traditional estimate of internal consistency. It reflects the average inter-correlation among items within a scale, assuming tau-equivalence. Values of .70 or higher are typically considered acceptable for research purposes, although moderately lower coefficients can be expected in short or heterogeneous scales. McDonald's omega (ω) provides a less restrictive estimate of internal consistency, accounting for unequal factor loadings among items. Unlike α, ω does not assume that all items contribute equally to the latent construct. Mean inter-item correlation (MIIC) represents the average of all pairwise correlations among items. It is a direct indicator of item homogeneity, with optimal values generally ranging from .20 to .50 depending on the breadth of the construct being measured. Lower MIIC values may reflect conceptual diversity among items, while excessively high values can suggest redundancy. Composite reliability (CR) was derived from CFA using standardised factor loadings and residual variances. CR reflects the proportion of variance in the observed variables that is explained by the underlying latent construct. Values above .60 are typically regarded as acceptable, indicating adequate construct reliability. Average variance extracted (AVE) was also estimated from CFA and represents the average proportion of variance that a construct explains in its indicators. AVE values above .50 indicate that the latent factor accounts for more than half of the variance in its indicators, supporting convergent validity.

Measurement of invariance testing. Measurement invariance was tested across groups defined by how frequently participants use LLMs to examine whether the factor structure of attitudes toward General and Primary LLMs is consistent across levels of user experience. Frequency of use is theoretically relevant because familiarity with LLMs reflects differential exposure to the technology, which may influence the attitudes, but should not alter the underlying construct being measured if the scale is valid.

We evaluated whether the factor structure underlying the acceptance and fear dimensions of AT-GLLM and AT-PLLM was equivalent for participants with lower versus higher levels of LLM experience.

A series of increasingly restrictive multi-group confirmatory factor analyses (CFA) was conducted using the robust maximum likelihood estimator (MLR) in lavaan. Models tested configural, metric, scalar, and strict invariance by sequentially constraining factor loadings, intercepts, and residual variances to equality across groups while retaining correlated factors. Model fit was evaluated using the comparative fit index (CFI), Tucker-Lewis index (TLI), root mean square error of approximation (RMSEA) with 90% confidence intervals, and standardised root mean square residual (SRMR). Evidence for invariance was determined using established practical cut-offs: a change of $\Delta CFI \leq 0.010$ and $\Delta RMSEA \leq 0.015$ between successive models was taken to indicate that the more constrained model did not significantly worsen fit. Where these criteria were met, the measurement model was considered invariant across levels of LLM use frequency.

External Validation. External validation was conducted using three theoretically relevant variables: self-efficacy, and the two components of ATAI (AI acceptance, and AI fear). Prior research suggests that higher self-efficacy is positively associated with the adoption and trust of AI systems (Naiseh et al., 2025) (Wang & Chuang, 2023). Individuals who perceive themselves as competent in using digital technologies tend to display more favorable attitudes toward AI applications and lower levels of anxiety. Accordingly, we hypothesized that there will be a positive relationship between self-efficacy and acceptance attitude toward LLMs and a negative relationship between self-efficacy and fear attitude toward LLMs. The ATAI scale was included as an external validation instrument to assess the convergent and discriminant validity of the new AT-GLLM and AT-PLLM measures. Given that ATAI served as the conceptual foundation for the new scales and ATAI and AT-GLLM measure general attitudes, stronger correlations were expected between ATAI and AT-GLLM than between ATAI and AT-PLLM.

To evaluate these relationships, a structural equation model (SEM) was estimated, including four dependent variables, AT-GLLM Acceptance, AT-GLLM Fear, AT-PLLM Acceptance, and AT-PLLM Fear, each regressed on three predictors: Self-efficacy, ATAI Acceptance, and ATAI Fear. The SEM was estimated using the lavaan package with the maximum likelihood method and NLMINB optimization. All variables were standardized prior to analysis. Nonparametric bootstrapping with 1,000 samples was applied to generate robust standard errors, and model fit was evaluated using CFI, TLI, RMSEA, and SRMR indices.

3. Results

3.1. Descriptive statistics and correlations

Table 2 and Table 3 presents the descriptive statistics and Pearson correlations for the study variables. The mean scores generally indicated high levels of acceptance and low levels of fear toward both AT-

GLLM and AT-PLLM. Similar pattern was observed for general AI acceptance and fear, confirming a consistent pattern of greater acceptance than fear across all levels of technology familiarity. The correlation results showed strong positive associations among the various acceptance dimensions, as well as strong positive associations among the fear dimensions. In contrast, acceptance measures were negatively correlated with fear measures, indicating that individuals who held more positive attitudes toward LLMs and AI tended to experience less apprehension toward them. Additionally, self-efficacy was positively correlated with acceptance and negatively correlated with fear, suggesting that individuals with higher confidence in using AI technologies were more likely to hold a positive attitude toward them and less likely to perceive them as threatening.

Table 2. Descriptive statistics between the study variables (N= 576).

	Descriptive statistics			
Scale	Mean	SD	Skewness	Kurtosis
1. AT-GLLM acceptance	15.30	2.85	-0.60	0.49
I trust them.	7.46	1.71	-0.89	1.58
They will benefit humankind.	7.84	1.48	-0.89	2.14
2. AT-GLLM fear	7.07	4.80	0.72	0.28
I fear them.	1.80	2.28	1.68	2.27
They will destroy humankind.	1.33	1.71	1.83	4.04
They will cause job losses.	3.93	2.57	0.20	-0.89
3. AT-PLLM acceptance	15.50	2.79	-0.66	0.58
I trust it.	7.56	1.67	-1.03	2.19
It will benefit humankind.	7.94	1.47	-0.53	0.39
4. AT-PLLM fear	6.99	4.72	0.81	0.77
I feat it.	1.88	2.27	1.67	2.50
It will destroy humankind.	1.18	1.63	1.88	4.52
It will cause job losses.	3.93	2.52	0.19	-0.83
5. ATAI acceptance	15.15	2.64	-0.70	1.19
I trust AI.	7.27	1.61	-0.84	1.49
AI will benefit humankind.	7.89	1.46	-1.18	3.69
6. ATAI fear	8.41	5.08	0.51	0.05
I fear AI.	2.47	2.15	1.06	0.91
AI will destroy humankind.	1.49	1.84	1.58	2.93
AI will cause many job losses.	4.45	2.47	-0.05	-0.81
7. Self-efficacy	7.33	1.45	0.07	-0.622

Table 3. Pearson correlations between the study variables (N= 576).

Scale	Correlations					
	1	2	3	4	5	6
1. AT-GLLM acceptance	-					
2. AT-GLLM fear	-.411***	-				
3. AT-PLLM acceptance	.828***	-.365***	-			
4. AT-PLLM fear	-.398***	.805***	-.417***	-		
5. ATAI acceptance	.808***	-.402***	.799***	-.413***	-	
6. ATAI fear	-.461***	.733***	-.460***	.782***	-.467***	-
7. Self-efficacy	.368***	-.227***	.331***	-.240***	.384***	-.254***

3.2. Confirmatory factor analysis

The two-factor models of the AT-GLLM and AT-PLLM were evaluated through confirmatory factor analyses (CFA) using the full dataset. Table 4 presents the model fit indices and standardized estimates for both scales

***AT-GLLM Scale*:** The analysis demonstrated an acceptable to good model fit: $\chi^2(3) = 16.870$, $p = .001$; CFI = .973; TLI = .910; RMSEA = .090; SRMR = .037. All standardized factor loadings were significant ($p < .001$), and ranged from .47 to .84, indicating that each variable contributed meaningfully to its latent construct (see Table 4). The latent factors showed a moderate negative correlation ($r = -.536$, $p < .001$), suggesting that greater fear were moderately associated with lower acceptance towards LLMs. Assessment of the model residuals revealed no evidence of local misfit, as all residual correlations being small in magnitude. The largest residual correlation (.097) was below the threshold of .10, indicating that the model adequately represented the item relationships while preserving local independence

***AT-PLLM Scale*:** The CFA results exhibited a very good model fit: $\chi^2(3) = 13.634$, $p = .003$; CFI = .979; TLI = .930; RMSEA = .078; SRMR = .031. All items loaded significantly and meaningfully onto their respective latent factors, with loadings ranging from .21 to .65 (see Table 4). The latent variables were negatively correlated ($r = -0.511$, $p < .001$), suggesting that lower fear toward primary LLMs was associated with higher acceptance. An inspection of the standardized residual covariance matrix revealed no substantial discrepancies between model and data, as all residual correlations were small in magnitude. The largest observed residual (.080), was below the .10 threshold, indicating minimal unexplained covariance.

Table 4. CFA factor loadings and R-squares

		AT-GLLM		AT-PLLM	
Factor	Item	Loading	R^2	Loading	R^2
Acceptance	I trust it/them	.844	. 712	.737	.544
	It/They will benefit humankind	.709	.502	.781	.610
Fear	I fear it/them	.466	.217	.457	.209
	It/They will destroy humankind	.743	.552	.807	.651
	It/They will cause job losses	.500	.250	.485	.235

Note: R^2 represents the proportion of variance in each item explained by the factor it loads on.

3.3. Reliability analysis

Reliability was evaluated for the Fear and Acceptance subscales of the ATAI, Personal AT LLM, and General AT LLMs questionnaires. The analyses used Cronbach's α, McDonald's ω (total), the mean inter-item correlation (MIIC), composite reliability (CR), and average variance extracted (AVE).

Across all scales, internal consistency ranged from modest to good (α = .54-.74; ω = .61-.75) (Table 5). The two-item Acceptance subscales showed the highest reliability (α = .64-.74), whereas the Fear subscales, each with three items, showed moderate consistency (α = .54-.68). MIIC values indicated acceptable inter-item homogeneity, particularly for the shorter Acceptance scales (.48-.59). Composite reliability (CR) and AVE followed the same pattern, with the highest values for the Acceptance subscales (CR = .48-.60; AVE = .48-.60).

Table 5. Reliability statistics for ATAI, Personal AT LLM, and General AT LLMs questionnaires

Questionnaire	Subscale	n_items	n	alpha	MIIC	omega_total	CR	AVE
ATAI	Fear	3	576	0.679	0.423	0.688	0.424	0.424
	Acceptance	2	576	0.642	0.475	0.647	0.479	0.479
Personal AT LLM	Fear	3	576	0.549	0.32	0.625	0.385	0.385
	Acceptance	2	576	0.728	0.577	0.732	0.578	0.578
General AT LLMs	Fear	3	576	0.541	0.308	0.608	0.367	0.367
	Acceptance	2	576	0.741	0.594	0.746	0.595	0.595

Note. α = Cronbach's alpha; MIIC = mean inter-item correlation; ω = McDonald's omega (total); CR = composite reliability; AVE = average variance extracted.

Taken together, internal consistency was highest for the Acceptance subscales, reflecting their focused content and strong inter-item associations. The Fear subscales demonstrated moderate reliability, consistent with their broader conceptual scope. All reliability indices were within acceptable limits for research use, and the pattern of results indicates that the Acceptance scales are psychometrically robust, whereas the Fear scales capture more heterogeneous attitudinal content.

3.4. Measurement invariance

AT-GLLM scale

Measurement invariance of the attitudes toward General LLMs (GLLM) was tested across two groups defined by LLM use frequency (Low = 423; High = 153). The configural model, representing equal factor structure but freely estimated parameters, achieved acceptable fit ($\chi^2(8) = 22.56$, $p = .004$; CFI = 0.940; TLI = 0.850; RMSEA = 0.079, 90% CI [0.041, 0.120]; SRMR = 0.034). This supports the presence of a similar latent structure for acceptance and fear across levels of LLM use. Imposing equality constraints on the factor loadings (metric model) produced comparable fit ($\chi^2(11) = 26.92$, $p = .005$; CFI = 0.934; TLI = 0.881; RMSEA = 0.071). The change relative to the configural model (ΔCFI = -0.006; ΔRMSEA = -0.009) fell well within recommended limits, and the Satorra-Bentler χ^2 difference test was non-significant ($\Delta\chi^2 = 3.49$, $\Delta df = 3$, $p = .322$), supporting metric invariance.

Adding intercept constraints (scalar model) led to a small further reduction in fit ($\chi^2(14) = 32.73$, $p = .003$; CFI = 0.923; TLI = 0.890; RMSEA = 0.068). However, the ΔCFI (–0.012) and ΔRMSEA (-0.003) again remained below conventional cut-offs (ΔCFI ≤ 0.010-0.020; ΔRMSEA ≤ 0.015), and the Satorra-Bentler difference ($\Delta\chi^2 = 4.48$, $\Delta df = 3$, $p = .214$) was not significant, indicating that intercept equality was tenable. Finally, constraining residual variances (strict model) improved fit markedly ($\chi^2(19) = 27.29$, $p = .098$; CFI = 0.966; TLI = 0.964; RMSEA = 0.039, 90% CI [0, 0.062]; SRMR = 0.063), with ΔCFI = +0.043 and ΔRMSEA = –0.029 relative to the scalar model. The scaled χ^2 difference was non-significant ($\Delta\chi^2 = 3.28$, $\Delta df = 5$, $p = .656$).

Taken together, the AT-GLLM model demonstrated full configural, metric, scalar, and strict invariance across lower and higher frequency LLM users. This indicates that the acceptance and fear dimensions of general LLM attitudes are measured equivalently across individuals with differing levels of LLM use experience.

AT-PLLM scale

Similar to AT-GLLM scale, measurement invariance of the attitudes toward Primary LLMs was tested across two groups defined by self-reported LLM use frequency (Low = 423; High = 153). The configural model provided a good fit to the data ($\chi^2(8) = 11.23$, $p = .189$; CFI = 0.986; TLI = 0.964; RMSEA = 0.037, 90% CI [0, 0.081]; SRMR = 0.024), supporting the equivalence of the factor structure across groups.

Constraining factor loadings to equality (metric model) resulted in a slight decrease in fit ($\chi^2(11) = 18.51$, $p = .071$; CFI = 0.967; RMSEA = 0.049), but the changes relative to the configural model (ΔCFI = –0.019; ΔRMSEA = +0.011) were within recommended thresholds (ΔCFI ≤ 0.010-0.020; ΔRMSEA ≤ 0.015). The Satorra–Bentler scaled χ^2 difference was non-significant ($\Delta\chi^2 = 3.59$, $\Delta df = 3$, $p = .309$), indicating that constraining loadings did not significantly worsen fit.

Adding intercept constraints (scalar model) slightly improved overall fit ($\chi^2(14) = 19.60$, $p = .143$; CFI = 0.975; RMSEA = 0.037), with ΔCFI = +0.008 and ΔRMSEA = -0.011 relative to the metric model. The Satorra–Bentler difference test was again non-significant ($\Delta\chi^2 = 2.60$, $\Delta df = 3$, $p = .457$). Constraining residual variances (strict model) produced excellent fit ($\chi^2(19) = 21.15$, $p = .328$; CFI = 0.991; RMSEA = 0.020), with ΔCFI = +0.015 and ΔRMSEA = -0.017 compared with the scalar model. The Satorra-Bentler scaled difference was non-significant ($\Delta\chi^2 = 3.83$, $\Delta df = 5$, $p = .574$).

Across all levels, fit indices remained well within accepted criteria, and no step produced a statistically significant deterioration in model fit. These results indicate that the PLLM measure demonstrates configural, metric, scalar, and strict invariance across lower and higher frequency LLM users, supporting equivalence of the factor structure, loadings, intercepts, and residual variances between groups.

3.4. External validation

This analysis examined the external validity of the AT-GLLM and AT-PLLM scales by assessing their relationships with self-efficacy (M = 7.33, SD = 1.45) and general attitudes toward AI (ATAI-acceptance: M = 15.30, SD = 2.85; ATAI-fear: M = 7.07, SD = 4.80). Using structural equation modeling (SEM), we tested whether self-efficacy and general attitudes toward AI predicted the four LLM-specific attitude dimensions: AT-GLLM-Acceptance, AT-GLLM-Fear, AT-PLLM-acceptance, and AT-PLLM-fear.

The model demonstrated a good fit: $\chi^2(3) = 31.586$, p = <.001; CFI = .990; TLI = .930; RMSEA = .129; SRMR = .079. The model accounted for a substantial proportion of variance across outcomes, explaining 61.5% of the variance in AT-GLLM Acceptance ($R^2 = .615$), 56.3 % in AT-GLLM Fear ($R^2 = .563$), 63.0% in AT-PLLM Acceptance ($R^2 = .630$), and 61.0% in AT-PLLM Fear ($R^2 = .610$). These results indicate that the external predictors (Self-efficacy, ATAI-Acceptance, and ATAI-Fear) collectively provided significant explanatory power for attitudes toward both general and primary LLMs.

For AT-GLLM acceptance, the SEM analysis showed that self-efficacy had a small to none impact on Accepting general LLMs (β = .034, SE = .031, 95% CI [-0.023, 0.094], p = .283). ATAI-Acceptance increased their likelihood of accepting general LLMs (β = .726, SE = .042, 95% CI [0.645, 0.806], p <.001), indicating a close alignment between the variables. ATAI fear had a small association with AT-GLLM Acceptance (β = -.112, SE = .031, 95% CI [-0.176, -0.051], p <.001), indicating less fear of AI was associated with more Acceptance of LLMs.

For AT-GLLM Fear, ATAI Fear emerged as a strong and positive predictor (β = .732, SE = .042, 95% CI [0.645, 0.809], p <.001), indicating a close alignment between both fear toward AI and general LLMs.

Neither ATAI-Acceptance (β = -.042, SE = .038, 95% CI [-0.122, 0.027], p = .261) nor Self-efficacy (β = -.038, SE = .029, 95% CI [-0.096, 0.023], p = .196) were significantly predictors of AT-GLLM Fear.

For AT-PLLM Acceptance, self-efficacy again had a non-significant effect on accepting primary LLMs (β = -.003, SE = .031, 95% CI [-0.064, 0.063], p = .925). ATAI-Acceptance showed a strong significant effect (β = .760, SE = .040, 95% CI [0.686, 0.840], p <.001) indicating that people with higher AI acceptance tend to also have greater acceptance of primary LLMs. ATAI Fear had a small association with AT-GLLM acceptance (β = -.084, SE = .031, 95% CI [-0.144, -0.023], p =.008), indicating less fear of AI was associated with more acceptance of primary LLMs.

For AT-PLLM Fear, ATAI Fear remained a strong positive predictor (β = .771, SE = .042, 95% CI [0.680, 0.848], p <.001), indicating that people who fear AI are more likely to fear primary LLMs. Neither ATAI-Acceptance (β = -.026, SE = .038, 95% CI [-0.111, 0.046], p = .493) nor Self-efficacy (β = -.018, SE = .030, 95% CI [-0.073, 0.043], p = .550) were significantly predictors of AT-GLLM Fear.

4. Discussion

This study aimed to adapt and validate two scales to measure toward LLMs for use within the Chinese context, which were originally developed in the UK population (Liebherr et al., 2025). The two scales are the Attitudes Toward General Large Language Models scale (AT-GLLM) and the Attitudes Toward Primary Large Language Model scale (ATPLLM). The AT-GLLM is designed to measure individuals' attitudes toward large language models in general and concerns about LLMs as a technology class. In contrast, the AT-PLLM focuses specifically on attitudes toward one's primary or most frequently used LLM. Building on previous validations in the UK, this study examined whether the factorial structure, reliability, and validity of the scales could be replicated among Chinese participants. Overall, the findings confirmed that both scales are psychometrically sound, conceptually stable, and suitable for assessing Chinese users' attitudes toward LLMs.

Compared to the validation of the English version of the AT-GLLM and AT-PLLM scales as well as the Arabic version (Barajeeh et al., 2025b), the internal reliability indices obtained for the Chinese sample were broadly consistent and within the expected range. In the Chinese data, Cronbach's alpha values for AT-GLLM were .54 for fear and .74 for acceptance, and for AT-PLLM were .55 for fear and .73 for acceptance, indicating moderate to good internal consistency across both scales. As in the UK (AT-GLLM α = .79 for acceptance and .78 for fear; AT-PLLM α = .75 and .74) and Arabic (AT-GLLM α = .68 for acceptance and .75 for fear; AT-PLLM α = .68 and .68) studies, the Acceptance subscales exhibited stronger internal reliability than the Fear subscales.

Composite reliability (CR) and average variance extracted (AVE) results supported this pattern, with CR values of .595 and .578 and AVE values of .595 and .578 for the General and Primary Acceptance subscales, respectively, indicating solid construct reliability and convergent validity. By contrast, the relatively lower AVE values for the Fear subscales (.367–.385) are consistent with findings from Sindermann et al. (2021), who reported a weak loading for the "job loss" item in the Chinese ATAI sample

(λ = .23) in the established ATAI measure. This suggests that Chinese respondents differentiate between pragmatic concerns (e.g., employment impact) and broader or ethical fears. Such conceptual diversity reduces shared variance among fear items, thereby lowering AVE, but it does not undermine the validity of the construct. This interpretation is supported by our CFA results, where the fear items “It/They will destroy humankind” showed the strongest factor loadings (λ ≈ .75–.81), while “It/They will cause job losses” and “I fear it/them” showed weaker associations (λ ≈ .46–.50). These differences reflect those observed in the established ATAI measure and suggest that Chinese respondents view existential risks as more central to AI-related fear than economic or personal concerns.

When testing the measurement invariance of the AT-GLLM and AT-PLLM scales, we compared individuals who used LLMs less often with those who engaged with them more frequently. This measurement invariance test allowed us to test weather the scale operate equivalently across groups that could differ meaningfully in their experience with the construct (Putnick & Bornstein, 2016). The split was based on the median value of self-reported use, which happened to fall at seven on the ten-point scale. As most participants reported use frequencies near the median value, the resulting groups different in size. The group distribution likely reflect broader trends in LLM engagement, where regular use has become increasingly common. The scales showed stable factor structures across both the larger group of moderate-to-frequent users and the smaller group of intensive users, indicating that people with very different levels of experience interpreted the underlying constructs of acceptance and fear in the same way. These findings support the stability and robustness of the scales across varying levels of user familiarity, suggesting that the AT-GLLM and AT-PLLM can be used confidently in future research with diverse samples.

The external validation analysis confirmed that the ATAI-acceptance dimension strongly predicted both acceptance toward general (AT-GLLM) and primary (AT-PLLM) large language models, while ATAI-fear was a strong predictor of corresponding LLM-specific fear. These findings align with the original UK and the Arabic validations of the AT-GLLM and AT-PLLM scales, confirming the theoretical continuity of the ATAI framework across cultural contexts and supporting the convergent validity of the Chinese versions of both scales. As in the Arabic validation, self-efficacy did not significantly predict either acceptance or fear toward LLMs once general AI attitudes were accounted for, contrasting with the UK sample where a small but significant effect emerged for AT-PLLM acceptance. This pattern suggests that in more collectivist societies, such as China and the Arab world, attitudes toward AI and LLMs may depend less on individual confidence in using technology and more on collective trust and perceived societal benefits. Empirical evidence from Chinese contexts supports this interpretation: studies have shown that AI self-efficacy influences AI acceptance primarily through trust in institutions and perceived moral alignment (Mao & Liu, 2025) (J. Li et al., 2021). For instance, Mao and Liu (2025) found that AI self-efficacy enhances AI acceptance mainly through trust, while Li et al. (2021) demonstrated that institutional commitment and leadership norms play stronger roles in shaping trust than AI self-efficacy.

Consistent with the English and Arabic versions, a substantial proportion of variance in LLM attitudes remained unexplained by general AI attitudes, ranging from approximately one-third to nearly half of the

total variance in prior studies. This residual variance reflects discriminant validity, suggesting that while attitudes toward LLMs share a strong conceptual foundation with general AI perceptions, they also capture distinct affective and ethical dimensions. These dimensions potentially stem from the interactive, conversational, and anthropomorphic nature of LLMs. Unlike traditional AI chatbots, LLMs invite social engagement through human like style conversation which can evoke empathy and relational bonding. Previous work showed that anthropomorphism increases trust toward AI agents (Natarajan & Gombolay, 2020). Studies have also shown that perceived ease of use, usefulness and social influence increase positive attitude toward LLM and its use (Abdaljaleel et al., 2024) (Liu et al., 2024). Taken together, our findings indicate that while the AT-LLM framework generalizes well cross-culturally, the psychological factors that could play role in the LLM acceptance differ culturally.

Beyond these findings, the study contributes to the going discussion regarding the structure of attitudes towards AI. While some models such as the General Attitudes Toward Artificial Intelligence Scale (GAAIS) (Schepman & Rodway, 2020) and the Attitudes Toward Artificial Intelligence (ATAI) (Sindermann et al., 2021), support a two-factor model, other instruments such as ATTARI-12 (Stein et al., 2024)and the AIAS-4 (Grassini, 2023) proposed a unidimensional construct. However, we argue that a single-factor solution may not adequately capture the complexity of people's responses to AI, as individuals can simultaneously fear and accept these technologies. For instance, one may be fascinated by the technological possibilities of LLMs while feeling uneasy about their ethical implications, a pattern described in the literature as technological ambivalence (Bala et al., 2017). The present findings support the two-factor approach and its validity across different cultures, reinforcing that acceptance and fear can coexist within individuals. The present findings therefore support a two-factor conceptualization and its cross-cultural validity, reinforcing that acceptance and fear can coexist within individuals.

Beyond its theoretical contribution, this study also carries practical implications for the responsible development and deployment of LLMs. As LLMs continue to evolve rapidly, understanding public and professional attitudes toward them is critical for ensuring reasonable, trustworthy, and socially aligned design and use. Insights from validated attitude measures such as the AT-GLLM and AT-PLLM can inform developers, policymakers, and educators about how people perceive the benefits and risks of these technologies, thereby guiding user-centered and ethically grounded innovation. Moreover, by providing validated instruments for the Chinese context, this study contributes to a more inclusive understanding of global AI perceptions. This is important because China has emerged as a major force in the global LLM ecosystem, particularly through the rapid growth of open-source models developed domestically (e.g., Qwen and DeepSeek), which together have accounted for a significant share of global AI usage in 2025 (M. Li, 2025) (Xu, 2025). As Chinese-developed LLMs continue to gain traction internationally, understanding domestic public attitudes becomes crucial for fostering trust, acceptance, and cross-cultural alignment in AI ecosystems. Attitude assessments can also help identify areas where public concern, such as ethics or privacy needs to be addressed, enabling developers and policymakers to tailor communication, governance, and design strategies that strengthen both public confidence and responsible technological progress.

The present study comes with the usual limitations of cross-sectional designs, which do not allow for causal inferences between variables. Second, the sample consisted primarily of Chinese online users, which may not fully represent the diversity of attitudes across different geographic regions, and levels of digital literacy within China. This limits the generalizability of the findings beyond comparable, digitally literate populations. Additionally, although the scales demonstrated sound psychometric properties, attitudes toward LLMs are likely to evolve rapidly alongside ongoing technological advancements, policy changes, and public discourse. Longitudinal research would therefore be valuable to assess how exposure, trust, and ethical perceptions develop over time. Future studies should also complement self-report data with behavioral indicators, for instance, examining whether more positive attitudes toward LLMs predict actual use frequency, reliance patterns, or integration into professional workflows.

Despite these limitations, the present findings indicate that the personal and general LLM scales are reliable and valid tools for assessing attitudes toward LLMs in the Chinese context. They can therefore be confidently used in future cross-cultural research and applications aimed at understanding the social and psychological dimensions of LLM adoption.

5. Conclusion

The present study provides evidence that the AT-GLLM and AT-PLLM as valid and reliable instruments for assessing acceptance and fear of LLMs within the Chinese context. Their stable factor structure, satisfactory reliability, and clear associations with general attitudes toward AI demonstrate that these scales capture meaningful psychological constructs. By offering culturally adapted and psychometrically sound tools, this work enables researchers and policymakers to better understand how users in China perceive and engage with LLMs. These measures can inform communication, governance, and design strategies that promote trust, ethical awareness, and responsible adoption.

Declarations

Acknowledgement. Open Access funding provided by the Qatar National Library. This publication was supported by NPRP 14 Cluster grant # NPRP 14C-0916–210015 from the Qatar National Research Fund (a member of Qatar Foundation). The findings herein reflect the work and are solely the responsibility of the authors.

Authors' Contribution. SA: Conceptualized the study, performed the analysis, curated the data, and wrote the first draft. **AY**: Conceptualized the study, performed the analysis, reviewed and edited the paper. **HY**: Translated into Chinese, curated the data, reviewed and edited the paper. **XW**: Contributed to the method, validated the analysis, reviewed and edited the paper. **JC**: Translated into Chinese, curated the data, reviewed and edited the paper. **TYM**: Translated into Chinese, curated the data, reviewed and edited the paper. **GX**: Reviewed and edited the paper. **CM**: Reviewed and edited the paper. **RA**: Conceptualized the study, validated the analysis, and reviewed and edited the paper.

Conflict of Interest: The authors declare no competing interests.

Ethical Approval and accordance

This study was approved by the Ethics Research Committee at Bournemouth University, UK (N62239, 03.03.2025) in accordance with the 1964 Helsinki Declaration.

Consent to participate

All participants provided informed consent prior to participation.

Consent to Publish

All participants provided informed consent for publication of their data prior to participation.

Data Availability and Materials: The questionnaires, dataset, and analysis files are available on the Open Science Framework (OSF) at the following link:

https://osf.io/qt8kf/overview?view_only=02a7a5a9076f4805a4c777341dae553c

References

1. Abdaljaleel, M., Barakat, M., Alsanafi, M., Salim, N. A., Abazid, H., Malaeb, D., Mohammed, A. H., Hassan, B. A. R., Wayyes, A. M., Farhan, S. S., Khatib, S. El, Rahal, M., Sahban, A., Abdelaziz, D. H., Mansour, N. O., AlZayer, R., Khalil, R., Fekih-Romdhane, F., Hallit, R., … Sallam, M. (2024). A multinational study on the factors influencing university students' attitudes and usage of ChatGPT. *Scientific Reports 2024 14:1*, *14*(1), 1983-. https://doi.org/10.1038/s41598-024-52549-8
2. Albuhairy, M. M., & Algaraady, J. (2025). Understanding User Perceptions of DeepSeek: A Mixed-Methods Sentiment and Thematic Analysis. *Authorea Preprints*. https://doi.org/10.36227/TECHRXIV.174140724.41500096/V1
3. Bala, H., Labonté-Lemoyne, E., & Léger, P. M. (2017). Neural Correlates of Technological Ambivalence: A Research Proposal. *Lecture Notes in Information Systems and Organisation*, *16*, 83–89. https://doi.org/10.1007/978-3-319-41402-7_11
4. Barajeeh, B., Yankouskaya, A., Alshakhsi, S., Maxwell Ho, C. S., Xu, G., & Ali, R. (2025a). Developing and Validating the Arabic Version of the Attitudes Toward Large Language Models Scale. *ArXiv Preprint*.
5. Barajeeh, B., Yankouskaya, A., Alshakhsi, S., Maxwell Ho, C. S., Xu, G., & Ali, R. (2025b). Developing and Validating the Arabic Version of the Attitudes Toward Large Language Models Scale. *ArXiv Preprint*.
6. Chang, W., Arcesati, R., & Hmaidi, A. (2025). *China's drive toward self-reliance in artificial intelligence: from chips to large language models | Merics*. https://merics.org/en/report/chinas-drive-toward-self-reliance-artificial-intelligence-chips-large-language-models
7. Choudhury, A., Shahsavar, Y., & Shamszare, H. (2025). User Intent to Use DeepSeek for Health Care Purposes and Their Trust in the Large Language Model: Multinational Survey Study. *JMIR Human*

Factors, *12*(1), e72867. https://doi.org/10.2196/72867

8. Di, W., Nie, Y., Chua, B. L., Chye, S., & Teo, T. (2023). Developing a Single-Item General Self-Efficacy Scale: An Initial Study. *Brief Article Journal of Psychoeducational Assessment*, *41*(5), 583–598. https://doi.org/10.1177/07342829231161884
9. Fried, A. J., Dorn, S. D., Leland, W. J., Mullen, E., Williams, D. M., Zaas, A. K., MacGuire, J., & Bynum, D. L. (2024). Large language models in internal medicine residency: current use and attitudes among internal medicine residents. *Discover Artificial Intelligence 2024 4:1*, *4*(1), 70-. https://doi.org/10.1007/S44163-024-00173-W
10. Grassini, S. (2023). Development and validation of the AI attitude scale (AIAS-4): a brief measure of general attitude toward artificial intelligence. *Frontiers in Psychology*, *14*, 1191628. https://doi.org/10.3389/FPSYG.2023.1191628/BIBTEX
11. Habeb Al-Obaydi, L., & Pikhart, M. (2025). DeepSeek as an AI-Powered Learning Tool: Uncovering EFL College Students' Preference and Satisfaction. *SAGE Open*, *15*(4). https://doi.org/10.1177/21582440251381683;PAGE:STRING:ARTICLE/CHAPTER
12. Li, J., Zhou, Y., Yao, J., & Liu, X. (2021). An empirical investigation of trust in AI in a Chinese petrochemical enterprise based on institutional theory. *Scientific Reports 2021 11:1*, *11*(1), 13564-. https://doi.org/10.1038/s41598-021-92904-7
13. Li, M. (2025). *USA vs China in AI & LLM: Statistics & Market Analysis [2025] | Second Talent*. Secondtalent. https://www.secondtalent.com/resources/usa-vs-china-ai-llm-statistics/
14. Lian, Y., Tang, H., Xiang, M., & Dong, X. (2024). Public attitudes and sentiments toward ChatGPT in China: A text mining analysis based on social media. *Technology in Society*, *76*, 102442. https://doi.org/10.1016/J.TECHSOC.2023.102442
15. Liebherr, M., Almourad, B., Alshakhsi, S. A., Montag, C., Almourad, M. B., Alshakshi, S., Xu, G., Ali, R., & Yankouskaya, A. (2025). *Developing and Validating the Attitudes Toward Large Language Models Scale*. https://doi.org/10.31234/osf.io/6yb5h_v2
16. Liu, F., Chang, X., Zhu, Q., Huang, Y., Li, Y., & Wang, H. (2024). Assessing clinical medicine students' acceptance of large language model: based on technology acceptance model. *BMC Medical Education 2024 24:1*, *24*(1), 1251-. https://doi.org/10.1186/S12909-024-06232-1
17. MacCallum, R. C., Widaman, K. F., Zhang, S., & Hong, S. (1999). Sample size in factor analysis. *Psychological Methods*, *4*(1), 84–99. https://doi.org/10.1037/1082-989X.4.1.84
18. Mao, Y., & Liu, Y. (2025). Examining of factors influencing AI acceptance: Self-efficacy, trust and hindrance Appraisal. *Proceedings of 2024 2nd International Conference on Information Education and Artificial Intelligence, ICIEAI 2024*, 763–769. https://doi.org/10.1145/3724504.3724630;PAGEGROUP:STRING:PUBLICATION
19. Montag, C., & Ali, R. (2023). Can we assess attitudes toward AI with single items? Associations with existing attitudes toward AI measures and trust in ChatGPT in two German speaking samples. *Accepted for Publication in Journal of Technology in Behavioral Science*. https://doi.org/10.21203/RS.3.RS-3325511/V1

20. Montag, C., Wang, Y., Ali, R., Elhai, J. D., & Yang, H. (2025). Shedding light on relationships between various attitudes towards AI measures and trust in generative AIs ChatGPT/Ernie Bot in a Chinese sample. *Journal of Psychology and AI*, *1*(1). https://doi.org/10.1080/29974100.2025.2503351
21. Naiseh, M., Babiker, A., Al-Shakhsi, S., Cemiloglu, D., Al-Thani, D., Montag, C., & Ali, R. (2025). Attitudes Towards AI: The Interplay of Self-Efficacy, Well-Being, and Competency. *Journal of Technology in Behavioral Science*, 1–14. https://doi.org/10.1007/S41347-025-00486-2
22. Natarajan, M., & Gombolay, M. (2020). Effects of anthropomorphism and accountability on trust in human robot interaction. *ACM/IEEE International Conference on Human-Robot Interaction*, 33–42. https://doi.org/10.1145/3319502.3374839;WGROUP:STRING:ACM
23. Niu, B., & Mvondo, G. F. N. (2024). I Am ChatGPT, the ultimate AI Chatbot! Investigating the determinants of users' loyalty and ethical usage concerns of ChatGPT. *Journal of Retailing and Consumer Services*, *76*, 103562. https://doi.org/10.1016/J.JRETCONSER.2023.103562
24. Pan, G., & Ni, J. (2024). A cross sectional investigation of ChatGPT-like large language models application among medical students in China. *BMC Medical Education*. https://doi.org/10.1186/s12909-024-05871-8
25. Putnick, D. L., & Bornstein, M. H. (2016). Measurement Invariance Conventions and Reporting: The State of the Art and Future Directions for Psychological Research. *Developmental Review : DR*, *41*, 71. https://doi.org/10.1016/J.DR.2016.06.004
26. Qu, S., Liu, L., Zhou, M., Zhou, C., & Campy, K. S. (2024). Factors influencing Chinese doctors to use medical large language models. *Digital Health*, *10*, 20552076241297236. https://doi.org/10.1177/20552076241297237
27. Rajik, J. A. (2024). ATTITUDES TOWARDS LARGE LANGUAGE MODELS AND MOTIVATIONS FOR THEIR USE: BASIS FOR CLASSROOM INTEGRATION. *European Journal of Education Studies*, *11*(12), 12–2024. https://doi.org/10.46827/EJES.V11I12.5727
28. Rosseel, Y. (2012). lavaan: An R Package for Structural Equation Modeling. *Journal of Statistical Software*, *48*, 1–36. https://doi.org/10.18637/JSS.V048.I02
29. Sarker, I. H. (2024). LLM potentiality and awareness: a position paper from the perspective of trustworthy and responsible AI modeling. *Discover Artificial Intelligence 2024 4:1*, *4*(1), 40-. https://doi.org/10.1007/S44163-024-00129-0
30. 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
31. Shahzad, M. F., Xu, S., & Javed, I. (2024). ChatGPT awareness, acceptance, and adoption in higher education: the role of trust as a cornerstone. *International Journal of Educational Technology in Higher Education 2024 21:1*, *21*(1), 1–23. https://doi.org/10.1186/S41239-024-00478-X
32. 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 - Kunstliche Intelligenz*, *35*(1), 109–118. https://doi.org/10.1007/S13218-020-00689-0/TABLES/5

33. Stein, J. P., Messingschlager, T., Gnambs, T., Hutmacher, F., & Appel, M. (2024). Attitudes towards AI: measurement and associations with personality. *Scientific Reports 2024 14:1*, *14*(1), 2909-. https://doi.org/10.1038/s41598-024-53335-2
34. Wang, Y. Y., & Chuang, Y. W. (2023). Artificial intelligence self-efficacy: Scale development and validation. *Education and Information Technologies 2023 29:4*, *29*(4), 4785–4808. https://doi.org/10.1007/S10639-023-12015-W
35. Wolf, E. J., Harrington, K. M., Clark, S. L., & Miller, M. W. (2013). Sample Size Requirements for Structural Equation Models: An Evaluation of Power, Bias, and Solution Propriety. *Educational and Psychological Measurement*, *73*(6), 913–934. https://doi.org/10.1177/0013164413495237;PAGE:STRING:ARTICLE/CHAPTER
36. Xu, E. (2025). *China's open-source models make up 30% of global AI usage, led by Qwen and DeepSeek | South China Morning Post*. Scmp. https://www.scmp.com/tech/tech-trends/article/3335602/chinas-open-source-models-make-30-global-ai-usage-led-qwen-and-deepseek?utm_source=chatgpt.com
37. Yeh, S. N., & Siah, C. J. R. (2025). Students' attitudes toward LLMs and its association with metacognitive abilities: A cross-sectional study. *Nurse Education in Practice*, *88*, 104567. https://doi.org/10.1016/J.NEPR.2025.104567
38. Zeng, L., Li, Q., Zuo, Y., Zhang, Y., & Li, Z. (2025). Perceptions and Attitudes of Chinese Oncologists Toward Endorsing AI-Driven Chatbots for Health Information Seeking Among Patients with Cancer: Phenomenological Qualitative Study. *Journal of Medical Internet Research*, *27*(1), e71418. https://doi.org/10.2196/71418
39. Zhao, Q., Chen, Y., & Liang, J. (2024). Attitudes and usage patterns of educators towards large language models: Implications for professional development and classroom innovation. *Academia Nexus Journal*, *3*(2).

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Appendix1.docx