

International Journal of Knowledge Processing Studies (KPS)



Homepage: <http://kps.artahub.ir/>



ORIGINAL RESEARCH ARTICLE

Subjective and Objective AI Literacy among Adult Residents of Tehran: Examining Their Correspondence

Zahra Ojagh^{1,*}, Mohammad Hasanzadeh²

¹ Associate Professor, Department of Communication of Science and Technology, Faculty of Cultural Studies and Communication, Institute for Humanities and Cultural Studies. zahraojagh1979@gmail.com, 0000-0001-8899-6495.

² Professor, Department of Information and Knowledge Science, Faculty of Management and Economics, Tarbiat Modares University, Tehran, Iran. hasanzadeh@modares.ac.ir, 0000-0002-6175-0855.

ARTICLE INFO

Article History:

Received: 2025-12-05

Revised: 2026-01-28

Accepted: 2026-02-27

Published Online: 2026-03-01

Keywords:

AI maturity, Subjective AI literacy, Objective AI literacy, Cognitive misalignment, Generational differences

Number of Reference: 20

Number of Figures: 0

Number of Tables: 4

DOI: 10.22034/kps.2026.592614.1289



ABSTRACT

This study examines AI maturity among citizens of Tehran by analyzing the gap between subjective and objective AI literacy as an indicator of cognitive misalignment in engagement with AI technologies. Using a quantitative cross-sectional survey design, data were collected from 507 residents of Tehran. Subjective AI literacy was measured using the AILS, while objective AI literacy was assessed through the AICOS, a specialized instrument comprising 51 objective and generative items. Reliability was confirmed using Cronbach's alpha and KR-20. Data were analyzed using independent-samples t tests, Mann-Whitney tests, Spearman correlation, and paired-samples t tests. The mean score for subjective AI literacy was 4.99 out of 7, whereas objective literacy scores across dimensions ranged from 0.07 to 0.10 on a 0-1 scale. A paired-samples t test confirmed a significant gap between the two dimensions with a very large effect size ($d_z = 4.27$). The correlation between subjective and objective AI literacy was weak and non-significant ($r_s = 0.062$). Generational analysis further showed that younger respondents (born after 1373) scored significantly higher in subjective AI literacy, but did not differ significantly from older respondents on most components of objective AI literacy. The findings indicate that AI maturity in this population is characterized by a structural misalignment between perceived competence and actual knowledge. Increased access to and use of AI tools has not necessarily translated into conceptual understanding, critical evaluation, or ethical awareness. Advancing AI maturity therefore requires educational and cultural policies that align use with understanding, critical judgment, and realistic self-assessment.

► **Citation:** Ojagh, Z., & Hassanzadeh, M. (2026). Subjective and Objective AI Literacy among Adult Residents of Tehran: Examining Their Correspondence. *International Journal of Knowledge Processing Studies (KPS)*, 6(1): 90-108. Doi: 10.22034/kps.2026.592614.1289

1. Introduction

Artificial intelligence (AI) is increasingly embedded in everyday communication, education, work, information seeking, and decision-making. As opportunities for direct interaction with AI systems have expanded, particularly through widely accessible generative AI applications, the ability to understand, evaluate, and use AI has become relevant beyond technical or professional settings. This broader societal engagement has increased interest in AI literacy, generally understood as a set of competencies that enable individuals to critically evaluate AI technologies, communicate and collaborate effectively with AI, and use AI as a tool in different contexts (Long & Magerko, 2020).

The expansion of AI use also raises an important measurement question. Individuals can be asked to evaluate their own AI-related knowledge and capabilities, or they can be assessed using tasks with objectively identifiable correct responses. These approaches do not provide identical forms of evidence. Self-report measures capture individuals' perceptions of their knowledge, skills, or capabilities, whereas objective assessments evaluate what respondents can demonstrate on the content being tested. Recent AI-literacy research has increasingly emphasized this distinction (Chiu et al., 2024), and evidence from studies employing both forms of assessment suggests that self-reported and objectively assessed AI literacy may show only limited correspondence (Zhang et al., 2026).

This distinction is important because conclusions about AI literacy may depend partly on how the construct is operationalized. A person may report confidence in using or evaluating AI without necessarily demonstrating corresponding knowledge on an objective assessment, while objective performance does not capture how capable individuals perceive themselves to be in interacting with AI. Neither form of assessment can therefore automatically be treated as a substitute for the other when the research question concerns both perceived capability and demonstrated performance.

A second relevant issue concerns variation across birth cohorts. Digital technologies are encountered within different historical and technological environments, and research has documented age- and generation-related differences in patterns of digital engagement. However, greater exposure to digital technologies should not be assumed to produce uniformly greater digital competence. Research challenging strong versions of the "digital native" assumption has demonstrated substantial variation in digital skills and practices even among younger populations (Hargittai, 2010; Calderón Gómez, 2019). Similarly, evidence of generational differences in engagement with emerging AI technologies does not establish corresponding differences in objectively demonstrated AI knowledge (Chan & Lee, 2023).

In Iran, AI development has increasingly become a subject of policy attention, including efforts to formulate policy approaches suited to the national context (Alinaghian et al., 2023). Against this background, citizen-level AI literacy represents a relevant area for empirical investigation. However, evidence concerning the relationship between subjective and objective AI-literacy assessments among heterogeneous adult community samples remains limited, particularly outside educational settings.

The present study addresses this issue by examining subjective and objective AI literacy among adults sampled from seven districts of Tehran. Subjective AI literacy is assessed through self-reported AI-related capabilities, whereas objective AI literacy is assessed through performance on an objective knowledge measure. Rather than assuming that the two measures represent interchangeable indicators or constructing a numerical "gap" between them, the study examines the extent to which they correspond empirically. It also investigates whether the two forms of assessment show similar or different patterns across two prespecified birth cohorts.

Accordingly, the study addresses the following research questions:

RQ1. What are the levels and distributions of subjective and objective AI literacy among adult respondents sampled from seven districts of Tehran?

RQ2. To what extent are subjective and objective AI literacy associated among these respondents?

RQ3. Do subjective and objective AI literacy differ between the prespecified earlier- and later-born cohorts?

2. Literature Review

2.1. Conceptualizing and Measuring AI Literacy

AI literacy has been conceptualized as extending beyond technical knowledge of algorithms or programming. Long and Magerko (2020) define AI literacy in terms of competencies that enable individuals to critically evaluate AI technologies, communicate and collaborate effectively with AI, and use AI as a tool in different settings. Subsequent research has operationalized AI literacy through a variety of instruments reflecting different conceptual domains, target populations, and measurement approaches.

An important distinction within this literature concerns subjective and objective assessment. Self-report instruments ask respondents to evaluate their own AI-related knowledge, capabilities, attitudes, or

behaviors. Such measures are appropriate when the construct of interest concerns perceived capability, confidence, or self-assessed competence. They do not, however, directly establish what respondents can demonstrate on a knowledge or performance assessment. Objective measures instead require respondents to answer questions or complete tasks for which performance can be evaluated against externally specified criteria. These approaches therefore provide different forms of evidence and should not be treated as interchangeable measures of demonstrated AI literacy (Chiu et al., 2024).

The growing AI-literacy measurement literature includes both self-report instruments, such as AILS, SNAIL, MAILS, AILQ, and PAILQ-6, and objective or performance-based assessments, including objective AI-literacy tests developed for educational populations, AICOS, and more recent generative-AI literacy assessments. These instruments differ substantially in their target populations, conceptual coverage, item formats, and intended interpretations. Consequently, scores produced by different AI-literacy instruments should not be assumed to represent equivalent measurement outcomes. Selected examples are summarized in Table 1.

Table 1. Selected Subjective and Objective Measures of AI Literacy

Instrument	Assessment Type	Items	Conceptual Domains / Structure	Response Format	Validation / Target Population	Source
Artificial Intelligence Literacy Scale (AILS)	Subjective	12	Awareness; Use; Evaluation; Ethics	7-point self-report	Adults / general users	Wang et al. (2023)
Scale for the Assessment of Non-Experts' AI Literacy (SNAIL)	Subjective	31	Technical Understanding; Critical Appraisal; Practical Application	Likert-type self-report	Non-expert adults	Laupichler et al. (2023)
Meta AI Literacy Scale (MAILS)	Subjective	34	Use & Apply AI; Understand AI; Detect AI; AI Ethics; Create AI as a separate construct; additional AI-related psychological competencies	Self-report	German-speaking adults	Carolus et al. (2023)
AI Literacy Questionnaire (AILQ)	Subjective	32	Affective; Behavioral; Cognitive; Ethical	5-point self-report	Secondary-school students	Ng et al. (2024)
Perceived Artificial Intelligence Literacy Questionnaire (PAILQ-6)	Subjective	6	Two-factor perceived AI-literacy structure	7-point self-report	UK adults	Grassini (2024)
AI Literacy Test	Objective	31 tasks/items	AI-literacy competencies based on a multidimensional competency framework	30 multiple-choice items and 1 sorting task	University students	Hornberger et al. (2023)
Objective AI Literacy Test	Objective	25	Objective assessment of AI-literacy competencies	Multiple-choice items	Grade 7–9 students	Chiu et al. (2024)
AI Competency	Objective	51 full; 18	Full version covers six	Objective-	Predominantly	Markus et

Objective Scale (AICOS)		short	AI-literacy subcompetencies; short version interpreted at the overall-score level	response items	non-student adult sample	al. (2025)
Generative AI Literacy Assessment Test (GLAT)	Objective / performance- based	20	Generative-AI literacy	Multiple- choice test	Higher- education students	Jin et al. (2025)

Note. The table presents selected instruments to illustrate variation in subjective and objective approaches to AI-literacy assessment and is not intended as an exhaustive inventory. “Subjective” refers to instruments based on respondents’ self-reports of their AI-related capabilities, perceptions, or competencies; “objective” refers to assessments scored against externally specified correct responses or performance criteria. The structures shown are those reported or used by the respective instrument developers and should not be interpreted as evidence that equivalent latent dimensions are measured across instruments. For AICOS, the full instrument contains 51 items, whereas the short version used in the present study contains 18 items and is analyzed only at the overall-score level.

2.2. Subjective and Objective AI-Literacy Assessment

Recent measurement research has increasingly emphasized the distinction between self-reported assessments of perceived AI-related capability and objective assessments of demonstrated AI-related knowledge or performance. Chiu et al. (2024), for example, explicitly distinguished perceived capability measured through self-report from objectively assessed AI literacy in developing an objective AI-literacy test. This distinction is methodologically important because a self-assessment and an objective test answer different questions: one concerns how individuals perceive their capabilities, whereas the other concerns their performance on the knowledge or competencies represented by the assessment.

Evidence from research employing both forms of assessment further indicates that the two should not be assumed to correspond closely. Zhang et al. (2026) reported misalignment between self-reported and objective-based AI-literacy measures, illustrating that higher self-assessment does not necessarily translate directly into higher objective performance. Such findings make

the empirical relationship between subjective and objective AI literacy a substantive research question rather than an assumption.

This distinction does not imply that self-report measures are methodologically inferior. When perceived capability, confidence, or self-assessed competence is itself the construct of interest, self-report is an appropriate measurement approach. Similarly, objective tests are appropriate when the research question concerns demonstrated knowledge or performance on specified content. The methodological problem arises when conclusions from one form of assessment are interpreted as though they necessarily represented the other.

Despite increasing attention to AI-literacy measurement, direct comparisons between subjective and objective assessments remain relatively limited, and much existing work has been conducted in educational populations. Consequently, the relationship between perceived AI-related capability and objectively assessed AI knowledge is less well characterized in heterogeneous adult community samples. The present study addresses this issue by examining both forms of assessment within the same adult sample.

2.3. Birth Cohort, Digital Experience, and AI Literacy

The rapid evolution of digital technologies has also generated interest in whether younger or more recently born individuals differ from older cohorts in their technological capabilities. Early “digital native” accounts proposed that individuals growing up with digital technologies might develop distinctive technological competencies. However, subsequent empirical research has challenged strong interpretations of this proposition.

Hargittai (2010), for example, demonstrated substantial variation in Internet skills and uses even among young adults commonly grouped within the “Net Generation.” Calderón Gómez (2019) similarly documented continuing digital

inequalities among young people, emphasizing differences in technological resources and practices rather than assuming uniform competence based on age or generation. Together, such evidence cautions against using birth cohort alone as an indicator of digital competence.

Emerging research suggests that generational differences may nevertheless occur in attitudes toward and engagement with AI technologies. Chan and Lee (2023), for example, identified differences between Gen Z students and Gen X and Millennial teachers in their experiences, perceptions, and intentions regarding generative AI. Such findings demonstrate that generational position may be associated with patterns of engagement with emerging AI technologies. They do not, however, establish that younger cohorts necessarily possess greater objectively demonstrated AI knowledge.

Birth cohort may therefore be relevant to AI-literacy research without being treated as a proxy for competence. In particular, cohort differences may appear differently depending on whether AI literacy is assessed through perceived capability or objective performance. Examining the two forms of assessment separately provides a means of determining whether any observed cohort pattern is consistent across measurement approaches.

Conceptual Framework

The conceptual framework of this study is organized around three related components: the multidimensional conceptualization of AI literacy, the distinction between subjective and objective forms of AI-literacy assessment, and birth cohort as a contextual grouping variable.

First, AI literacy is treated as a broad set of competencies rather than a single narrowly technical skill. Long and Magerko (2020) conceptualize AI literacy as encompassing competencies relevant to critically evaluating AI, communicating and collaborating with AI, and using AI effectively in different settings. Measurement instruments operationalize different portions of this broader conceptual domain. Consequently, conclusions about AI literacy must be

interpreted in relation to the specific form of evidence generated by the instrument used.

Second, the framework distinguishes subjective AI literacy from objective AI literacy at the level of measurement. Subjective assessment captures respondents' perceptions of their own AI-related capabilities. In the present study, this is operationalized using AILS, which assesses self-reported awareness, use, evaluation, and ethics (Wang et al., 2023). Objective assessment instead concerns demonstrated performance on items with identifiable correct responses. In this study, it is operationalized using the overall correctness score from the 18-item short version of AICOS (Markus et al., 2025). Consistent with recent measurement research (Chiu et al., 2024; Zhang et al., 2026), these measures are treated as related but non-equivalent forms of evidence.

The framework therefore does not assume that subjective and objective scores should be numerically equal, nor does it define the difference between their raw scores as an AI-literacy "gap." Because the instruments differ in response format, score range, construct coverage, and measurement properties, direct subtraction would not yield a defensible measure of calibration. Instead, correspondence is operationalized as the statistical association between subjective and objective scores. A stronger positive association would indicate greater correspondence in respondents' relative standing across the two measures, whereas a weak association would indicate limited correspondence. This approach does not, by itself, identify overestimation, underestimation, or metacognitive calibration.

Third, birth cohort is incorporated as a contextual grouping variable, rather than as an assumed determinant of AI literacy. Digital-divide research has demonstrated that access to and engagement with digital technologies are socially differentiated and that exposure alone does not guarantee equivalent skills or outcomes (Van Deursen & Van Dijk, 2011; Van Dijk, 2012, 2017). Similarly, critiques of strong "digital native" assumptions demonstrate that younger age or

greater exposure to technology cannot automatically be equated with greater competence (Hargittai, 2010; Calderón Gómez, 2019).

Accordingly, the framework does not predict that later-born respondents must exhibit higher AI literacy. Instead, it treats birth cohort as a basis for examining whether subjective and objective assessments display similar or different empirical patterns across the prespecified earlier- and later-born groups. This distinction is important because cohort differences in perceived capability need not be accompanied by corresponding differences in objectively assessed knowledge.

The resulting conceptual model therefore contains two primary measured constructs—subjective AI literacy and objective AI literacy—together with birth cohort as a grouping variable. The empirical analysis addresses three questions derived from this framework: the distributions of the two forms of AI-literacy assessment, the degree of correspondence between them, and whether each differs across the prespecified birth cohorts. Continuous birth year is additionally examined as a supplementary analysis rather than as a separate theoretical construct or research question.

3. Method

3.1. Study Design

This study employed a quantitative, cross-sectional survey design to examine subjective and objective AI literacy among adult respondents sampled from seven municipal districts of Tehran and to investigate the correspondence between these two forms of assessment. Subjective AI literacy was assessed using the Artificial Intelligence Literacy Scale (AILS; Wang et al., 2023), whereas objective AI literacy was assessed using the 18-item short version of the Artificial Intelligence Competency Objective Scale (AICOS; Markus et al., 2025).

The two instruments were treated as providing related but distinct forms of evidence about AI literacy. Accordingly, subjective and objective AI literacy were analyzed separately, and their

correspondence was examined through statistical association rather than by subtracting, equating, or transforming their scores onto an assumed common metric. No composite AI-maturity score or subjective–objective discrepancy index was constructed.

3.2. Sampling Design and Participant Recruitment

An initial sample-size planning benchmark of 384 respondents was obtained using Cochran's large-population formula under conventional assumptions of 95% confidence, 5% nominal precision, and maximum variability. Because individual respondents were ultimately recruited through a nonprobability procedure, these parameters are reported only as planning assumptions and should not be interpreted as design-based confidence or margin-of-error guarantees for the achieved sample. To accommodate the geographically structured design, the prespecified birth-year stratification, and potential data loss, the operational recruitment target was increased to 448 respondents.

To achieve geographic coverage across Tehran, the city's 22 municipal districts were first grouped into five geographic clusters: North, South, Central, East, and West. Districts were then randomly selected within each cluster, with one district selected from clusters containing two or four districts and two districts selected from clusters containing six or eight districts. This procedure resulted in the selection of seven municipal districts—Districts 2, 4, 5, 7, 10, 14, and 15. The planned allocation was 64 respondents per selected district, yielding a total target sample of 448 respondents.

The sampling plan additionally incorporated a prespecified birth-year stratification boundary at Solar Hijri year 1373 (approximately March 1994–March 1995). Within each selected district, the intended allocation was 32 respondents born before 1373 and 32 respondents born in 1373 or later. This boundary served as an operational sampling stratification criterion and was not intended to represent a universal definition of Generation Z or any other named generation.

Within the selected districts, fieldworkers approached potential participants and screened them for residence in the relevant district and the applicable birth-year stratum. Individual respondents were recruited through a quota-based intercept procedure rather than probability-based random selection from a complete population sampling frame. Residence in one of the seven selected districts was an intended eligibility requirement.

A total of 507 questionnaires were collected. All individuals who agreed to participate completed the questionnaire. However, because the total number of individuals approached during field recruitment was not systematically recorded, a conventional survey response rate could not be calculated.

For the present analyses, eligibility was restricted to respondents whose adult status during the data-collection period could be established and who met the prespecified district-residence criterion. Respondents born in Solar Hijri year 1387 or later ($n = 40$) were excluded because they were younger than 18 years during the data-collection period. A further 23 respondents born in 1386 were excluded because only year, rather than exact date, of birth had been recorded; consequently, their adult status at the time of participation could not be established with certainty. These exclusions yielded an adult pool of 444 respondents.

Of these 444 adults, 44 did not reside in one of the seven prespecified districts. These questionnaires had been completed voluntarily by individuals encountered during fieldwork who wished to participate despite not meeting the intended district-residence criterion. Because these cases represented deviations from the prespecified geographic eligibility criterion, they were excluded from the primary analysis irrespective of their responses to the AI-literacy measures.

The resulting primary analytic sample comprised 400 adult respondents from the seven selected districts: 59 from District 2, 60 from District 4, 60 from District 5, 59 from District 7, 59 from District 10, 59 from District 14, and 44 from District 15.

Although districts were randomly selected within the geographically structured sampling design, individual respondents within the selected districts were recruited through field-based quota/intercept procedures rather than probability sampling from a complete population sampling frame. In addition, the achieved district allocations differed from the planned equal allocations. Individual selection probabilities are therefore unknown, and the resulting sample should not be interpreted as a probability-representative sample of the adult population of Tehran or Iran. Accordingly, population-level prevalence estimates are not inferred from these data.

3.3. Measures

3.3.1. Subjective AI Literacy

Subjective AI literacy was assessed using the 12-item Artificial Intelligence Literacy Scale (AILS; Wang et al., 2023). AILS is a self-report instrument comprising four dimensions—awareness, use, evaluation, and ethics—with three items representing each dimension. Responses were recorded on a seven-point Likert-type scale.

Items 2, 5, and 11 are negatively worded and were reverse-scored before scale scores were calculated. Higher scores therefore represented higher self-reported AI literacy. The overall AILS score was calculated across the 12 items, and dimension scores were calculated for awareness, use, evaluation, and ethics in accordance with the original scale structure. Because AILS assesses respondents' perceptions of their AI-related capabilities, its scores are referred to throughout this study as subjective AI literacy rather than objectively demonstrated AI competence.

In the primary analytic sample, the overall 12-item AILS yielded Cronbach's $\alpha = .779$. This coefficient is interpreted as an indicator of the internal consistency of scores obtained in the present sample and not as evidence of construct validity or cross-cultural measurement equivalence.

3.3.2. Objective AI Literacy

Objective AI literacy was assessed using the 18-item short version of the Artificial Intelligence Competency Objective Scale (AICOS; Markus et al., 2025). The items require respondents to answer questions for

which a correct response can be identified. Each correct response was assigned a score of 1 and each non-correct response a score of 0, yielding an overall correctness score ranging from 0 to 18, with higher scores indicating a greater number of correctly answered items.

The 18-item short version was analyzed exclusively at the overall-score level. Dimension-specific objective literacy scores were not calculated because the scoring guidance for the short version supports interpretation at the overall AI-literacy level rather than extraction of separate subscale scores. Accordingly, although the full AICOS represents multiple AI-literacy subcompetencies, the present study makes no dimension-specific inferences from the 18-item version.

For the present study, the researchers added an "I do not know" response option to the AICOS items. This option was not part of the original response format and therefore represents a modification of the administered instrument. It was introduced to allow respondents to explicitly indicate uncertainty or lack of knowledge rather than requiring them to select one of the substantive response alternatives. For calculation of the overall correctness score, correct responses were scored 1, whereas both incorrect responses and "I do not know" responses were scored 0. "I do not know" responses were nevertheless retained as a distinct response category for descriptive analyses and were not coded as missing. Genuinely unanswered items were treated as missing.

No psychological or metacognitive interpretation was assigned to these response categories. In particular, selection of an incorrect substantive response was not interpreted as evidence of inattentiveness, and selection of "I do not know" was not treated as direct evidence of metacognitive awareness or accurate self-assessment. Moreover, because the "I do not know" option modified the original response format, its inclusion may have influenced response behavior and limits direct comparability with administrations using the original AICOS response format.

In the primary analytic sample, the dichotomously scored 18-item AICOS total yielded $KR-20 = .876$. This coefficient describes the internal consistency of the correctness scores in the present sample and does not, by itself, establish construct validity, cross-cultural equivalence, or adequate discrimination across the full range of AI-literacy ability.

3.4. Translation and Cultural Adaptation

The research team contacted the developers of both instruments by email, explained the purpose and context of the study, and requested access to the instrument materials. The developers provided the instruments for use in the study.

The English-language materials were initially translated into Persian by the principal researcher. The initial Persian translations were subsequently reviewed during two expert-panel sessions involving five specialists with backgrounds in communication, media management, Persian language and literature, systems design engineering, and information and knowledge science. The panel members were proficient in Persian and English, and each session lasted approximately 2.5 hours. Item wording and intended conceptual meaning were discussed, and revisions were made through panel discussion and agreement.

The resulting Persian versions were subsequently reviewed against the source materials in a separate face-to-face session with an additional specialist in computer science and engineering. Further revisions were made where necessary to improve linguistic clarity and conceptual appropriateness.

A back-translation procedure was then undertaken. The revised Persian versions were independently translated back into English by a specialist proficient in both Persian and English. Another specialist subsequently compared the original English materials with the back-translated versions for conceptual and semantic correspondence, and further revisions were made where necessary.

The resulting Persian questionnaire was pilot-tested with 10 participants drawn from student and employed populations.

Participants were asked to identify wording or expressions that they found difficult to understand, ambiguous, or insufficiently clear. Their feedback informed further editorial revisions before the main field administration.

These procedures were undertaken to support the linguistic clarity and face and content appropriateness of the Persian questionnaire. They should not be interpreted as establishing full psychometric or cross-cultural validation of Persian versions of either instrument. In particular, measurement invariance, criterion validity, and the factor structure of the translated instruments were not independently established in the present study. The adaptation procedures therefore provide evidence regarding the translation and review process rather than evidence of psychometric equivalence between the Persian and source-language versions.

This limitation is particularly relevant to AICOS because the source documentation indicates that its original validation was conducted using the German-language items, while the English-language wording supplied with the instrument served as recommended translations. The Persian version used in the present study was therefore subjected to the translation and review procedures described above but should not be regarded as an independently validated Persian version of AICOS.

3.5. Preliminary Reliability Assessment

Before the main field administration, internal consistency was examined in preliminary samples. The objective AI-literacy instrument was administered to 50 participants and yielded KR-20 = .767. The subjective AI-literacy instrument was administered to a separate sample of 70 participants and yielded Cronbach's α = .746.

KR-20 was used for the objective correctness score because the items were scored dichotomously, whereas Cronbach's alpha was used for the Likert-type subjective measure. These preliminary coefficients were treated as descriptive indicators of score consistency before the main administration rather than as evidence of construct validity or cross-cultural validity,

and no claim of psychometric adequacy was based solely on exceeding a fixed reliability threshold.

Internal consistency was subsequently re-estimated in the primary analytic sample. The overall AILS yielded α = .779, whereas the AICOS correctness score yielded KR-20 = .876. These coefficients are interpreted only as indicators of internal consistency of the observed scores in the present sample.

3.6. Questionnaire Administration and Standardization

The final questionnaire included the two AI-literacy instruments together with demographic and background questions. Respondents were also asked whether they had previously participated in AI-related training and about their use of ChatGPT.

Two written protocols were developed to standardize procedures during questionnaire preparation and piloting and during the main field administration. The field protocol specified procedures to be followed by fieldworkers during recruitment and questionnaire administration. A glossary was also prepared for terms identified during the adaptation process as potentially unfamiliar or difficult for respondents to understand. Fieldworkers could refer to this glossary when clarification was requested.

The purpose of the glossary was to standardize the intended meaning of potentially unfamiliar terminology rather than to indicate or facilitate a correct response to an objective test item. Nevertheless, because clarification of terminology during an objective knowledge assessment may potentially affect item functioning, the availability of such clarification is considered part of the administration procedure and is taken into account when interpreting the objective scores.

3.7. Ethical Considerations and Adult Eligibility

Participation was voluntary, and informed consent was obtained before questionnaire completion. Respondents were informed of the purpose of the research and the confidentiality of their responses.

Because exact dates of birth were not recorded, adult status could not be

established with certainty for respondents born in Solar Hijri year 1386. The primary analyses were therefore restricted to respondents whose adult status during the data-collection period could be established from the available birth-year information. Respondents born in 1387 or later and respondents born in 1386 whose adult status could not be established with certainty were excluded from the analyses, as described in Section 3.2.

3.8. Birth-Year Classification

Birth year was recorded using the Solar Hijri calendar. The sampling design used Solar Hijri year 1373 (approximately March 1994–March 1995) as a prespecified stratification boundary. Respondents born before 1373 were classified as the earlier-born cohort, whereas respondents born in 1373 or later were classified as the *later-born cohort*.

These categories are operational groupings specific to the sampling design and are not treated as conventional sociological generations. In particular, the 1373 boundary should not be interpreted as defining Generation Z, Millennials, Baby Boomers, or any other named generation.

Birth year was additionally retained as a continuous variable for supplementary analysis. This analysis provides information without imposing the 1373 dichotomization. However, because the study is cross-sectional, associations involving birth year cannot distinguish age-related, birth-cohort, and period-related processes and are therefore not interpreted causally.

3.9. Missing Data and Analytic Samples

No missing values were imputed. Analyses were based on complete observations for the variables required by each statistical procedure. Of the 400 respondents in the primary analytic sample, 395 had complete data for calculation of the overall AILS score, 393 had complete data for calculation of the 18-item AICOS correctness score, and 389 had complete data for both overall measures.

For AICOS, the researcher-added “I do not know” option was treated as an observed substantive response and scored as non-correct (0), rather than being treated as

missing. Only genuinely unanswered items were coded as missing.

Because the number of complete observations differed slightly across measures and analyses, the analytic sample size is reported for each principal statistical analysis rather than assuming a common N of 400 for all analyses.

3.10. Statistical Analysis

Descriptive statistics were calculated separately for subjective and objective AI literacy. For AILS, means, standard deviations, medians, skewness, and kurtosis were calculated for the overall score and its four established dimensions. For AICOS, descriptive analyses were restricted to the overall 0–18 correctness score. Correct, incorrect, and “I do not know” responses were additionally summarized descriptively to characterize the observed response pattern, including the concentration of scores at the lower end of the objective score distribution.

The correspondence between overall subjective and objective AI literacy was examined using Spearman’s rank-order correlation coefficient (r_s). Spearman correlation was selected because the objective score exhibited a pronounced floor concentration. The analysis assessed the monotonic association between the two measures rather than agreement or numerical equivalence. Because AILS and AICOS differ in construct operationalization, response format, and score range, no paired-samples mean comparison and no arithmetic subjective–objective discrepancy score were calculated.

For the prespecified birth-cohort comparison, overall subjective AI literacy was treated as the primary subjective outcome. The earlier- and later-born cohorts were compared using Welch’s independent-samples t test, which does not require the assumption of equal population variances. The four AILS dimension scores were examined as secondary outcomes. To control family-wise error across these four dimension-level comparisons, their p values were adjusted using Holm’s sequential procedure. The overall AILS comparison was not included in this multiplicity family

because it represented the prespecified overall subjective outcome rather than one of the four secondary dimension-level comparisons. Group means and standard deviations, 95% confidence intervals for mean differences, and Hedges' g were reported to characterize the magnitude and precision of observed differences.

Because the AICOS total score exhibited a pronounced floor concentration and extensive ties, the difference in objective AI literacy between the two birth cohorts was examined using the Mann–Whitney U test. Rank-biserial correlation was reported as the corresponding effect-size measure. Given the distributional characteristics of the objective score, the Mann–Whitney U test was interpreted as a rank-based comparison of the two group distributions and not specifically as a test of equality of medians.

As a supplementary analysis, Spearman correlations were calculated between continuous birth year and (a) overall subjective AI literacy and (b) overall objective AI literacy. These analyses examined the pattern of association without imposing the 1373 dichotomization. They were not interpreted as causal estimates of age or cohort effects.

All statistical tests were two-sided, with $\alpha = .05$ used as the nominal significance threshold. Exact p values were reported where informative, and inferential findings were interpreted together with effect-size estimates and 95% confidence intervals where applicable. Statistical analyses were conducted using IBM SPSS Statistics, version 26.

4. Findings

The primary analytic sample comprised 400 adult respondents who met both the adult-eligibility criterion and the prespecified district-residence criterion. Participants resided in the seven Tehran districts included in the sampling design: Districts 2 ($n = 59$), 4 ($n = 60$), 5 ($n = 60$), 7 ($n = 59$), 10 ($n = 59$), 14 ($n = 59$), and 15 ($n = 44$). Of the final sample, 229 respondents (57.3%) were women, 168 (42.0%) were men, and 3 (0.8%) preferred not to report their gender.

Using the birth-year boundary prespecified in the sampling design, 229 respondents (57.3%) were born before Solar Hijri year 1373 and 171 (42.8%) were born in 1373 or later. Solar Hijri year 1373 corresponds approximately to March 1994–March 1995. Because this boundary was used as a sampling stratification criterion rather than as a universal definition of a named generation, the two groups are referred to as the earlier-born and later-born cohorts.

Item-level missingness was limited but resulted in small variations in the analytic sample size across measures and analyses. No missing values were imputed; consequently, the analytic sample size is reported for each analysis.

4.1. Descriptive Results for Subjective and Objective AI Literacy

Subjective AI literacy was assessed using the AILS. Among respondents with complete overall AILS data ($n = 395$), the mean score was 4.98 on the 1–7 scale ($SD = 0.89$), and the median was 5.08. At the subscale level, the highest mean was observed for ethics ($M = 5.30$, $SD = 1.27$), followed by evaluation ($M = 5.13$, $SD = 1.00$), awareness ($M = 4.77$, $SD = 1.36$), and use ($M = 4.71$, $SD = 1.39$). The AILS distributions were negatively skewed, although the magnitude of skewness was modest. The overall 12-item AILS yielded Cronbach's $\alpha = .779$.

Objective AI literacy was assessed using the overall 18-item AICOS short-form correctness score. Dimension-specific scores were not calculated or interpreted. Among respondents with complete AICOS data ($n = 393$), the mean number of correct responses was 1.31 out of 18 ($SD = 2.64$), corresponding to an average of approximately 7.3% of the 18 items answered correctly per respondent. The median was 0, and observed scores ranged from 0 to 16. The distribution was strongly positively skewed (skewness = 2.71) and leptokurtic (excess kurtosis = 8.93), demonstrating a pronounced concentration of scores at the lower boundary. The AICOS correctness score yielded KR-20 = .876.

Among the 393 respondents with complete AICOS scores, 275 (70.0%) obtained no correct answers, and 258 (65.6%) selected “I

do not know” for all 18 items. Across all available AICOS item responses in the primary sample, 541 responses (7.5%) were correct, 1,187 (16.5%) were incorrect, and 5,454 (75.9%) were “I do not know.” These

item-response frequencies are reported descriptively and were not used to classify respondents into cognitive, metacognitive, or latent categories.

Table 2. Descriptive Statistics for Subjective and Objective AI Literacy

Measure	n	M	SD	Median	Skewness	Excess Kurtosis
AILS – Awareness	397	4.77	1.36	5.00	–0.63	–0.02
AILS – Use	399	4.71	1.39	5.00	–0.56	–0.23
AILS – Evaluation	398	5.13	1.00	5.33	–0.53	1.16
AILS – Ethics	400	5.30	1.27	5.33	–0.72	0.09
AILS – Overall	395	4.98	0.89	5.08	–0.25	–0.07
AICOS – Overall	393	1.31	2.64	0.00	2.71	8.93

Note. AILS scores range from 1 to 7, with higher scores indicating higher self-assessed AI literacy. The AICOS short-form overall score ranges from 0 to 18 and represents the number of correct responses. Sample sizes vary slightly because of item-level missingness. AICOS dimension-level scores were not calculated because the short form was analyzed only at the overall-score level.

5.2. Correspondence between Subjective and Objective AI Literacy

The association between subjective and objective AI literacy was examined among respondents with complete scores on both measures ($n = 389$) using Spearman's rank-order correlation.

Subjective and objective AI literacy were positively associated, $r_s = .161$, $p = .001$. The magnitude of this association was small, indicating limited correspondence between the two measures: respondents reporting higher subjective AI literacy tended, on average, to obtain somewhat higher objective scores, but the relationship between the two measures was weak.

No direct subtraction of AILS and AICOS scores was performed because the two instruments use different measurement metrics. Accordingly, no paired-samples comparison or individual discrepancy (“gap”) score was used to infer overestimation or underestimation.

4.3. Birth-Cohort Differences in Subjective and Objective AI Literacy

Birth-cohort differences were first examined for subjective AI literacy. The overall AILS score was treated as the primary subjective-literacy outcome. The

four subscale comparisons were treated as secondary analyses, with Holm adjustment applied across these four tests.

The later-born cohort reported higher overall subjective AI literacy ($M = 5.28$, $SD = 0.79$, $n = 169$) than the earlier-born cohort ($M = 4.75$, $SD = 0.89$, $n = 226$). The difference was statistically significant, Welch's $t(382.75) = -6.23$, $p < .001$, with Hedges' $g = -0.62$. The 95% confidence interval for the earlier-minus-later mean difference was $[-0.69, -0.36]$.

Secondary analyses showed higher subjective scores among later-born respondents for awareness, use, and evaluation. These differences remained statistically significant after Holm adjustment (all adjusted $p < .001$). In contrast, ethics scores did not differ significantly between the two cohorts (adjusted $p = .349$).

For objective AI literacy, the earlier-born cohort had a mean AICOS score of 1.39 ($SD = 2.80$, $Mdn = 0$; $n = 224$), compared with 1.21 ($SD = 2.40$, $Mdn = 0$; $n = 169$) in the later-born cohort. The Mann–Whitney U test indicated no statistically significant difference between the two cohort distributions, $U = 18,817.5$, $p = .903$, with an effect size close to zero ($r_{rb} = -.006$).

Thus, the observed birth-cohort pattern differed across the two forms of assessment: later-born respondents reported higher subjective AI literacy overall and on awareness, use, and evaluation, whereas no statistically detectable cohort difference was observed in overall objective AI literacy.

Table 3. Birth-Cohort Differences in Subjective and Objective AI Literacy

Outcome	Earlier-born <i>n</i>	Earlier-born <i>M</i> (<i>SD</i>)	Later-born <i>n</i>	Later-born <i>M</i> (<i>SD</i>)	Test statistic	95% CI	Effect size	<i>p</i> / Adjusted <i>p</i>
AILS – Awareness	227	4.43 (1.41)	170	5.22 (1.15)	$t(392.31) = -6.15$	$[-1.04, -0.54]$	$g = -0.60$	<.001*
AILS – Use	228	4.28 (1.42)	171	5.27 (1.13)	$t(395.51) = -7.70$	$[-1.24, -0.74]$	$g = -0.75$	<.001*
AILS – Evaluation	228	4.91 (1.02)	170	5.43 (0.89)	$t(386.68) = -5.43$	$[-0.71, -0.33]$	$g = -0.54$	<.001*
AILS – Ethics	229	5.35 (1.18)	171	5.23 (1.39)	$t(332.09) = 0.94$	$[-0.14, 0.38]$	$g = 0.10$	.349*
AILS – Overall	226	4.75 (0.89)	169	5.28 (0.79)	$t(382.75) = -6.23$	$[-0.69, -0.36]$	$g = -0.62$	<.001
AICOS – Overall	224	1.39 (2.80)	169	1.21 (2.40)	$U = 18,817.5$	—	$r_{rb} = -.006$	.903

Note. Earlier-born = born before Solar Hijri year 1373; later-born = born in 1373 or later. Solar Hijri year 1373 corresponds approximately to March 1994–March 1995. Welch's *t* tests were used for AILS comparisons. Hedges' *g* is reported as the standardized effect size; negative values indicate higher scores in the later-born group because differences were calculated as earlier-born minus later-born. Asterisks identify the four AILS subscale *p* values adjusted using the Holm procedure. The overall AILS comparison was specified separately and was not included in this multiplicity family. AICOS was compared using the Mann–Whitney *U* test.

4.4. Supplementary and Sensitivity Analyses

Because categorizing birth year reduces information and the prespecified Solar Hijri year 1373 boundary should not be interpreted as a universal generational threshold, supplementary analyses retained birth year as a continuous variable.

Table 4. Associations between Subjective AI Literacy, Objective AI Literacy, and Birth Year

Association	<i>n</i>	Spearman's <i>rs</i>	<i>p</i>
Subjective AI literacy × Objective AI literacy	389	.161	.001
Birth year × Subjective AI literacy	395	.365	<.001
Birth year × Objective AI literacy	393	.020	.690

Note. Higher Solar Hijri birth years indicate more recent birth. Spearman's rank-order correlations were used for these analyses. The continuous birth-year analyses were supplementary to the prespecified categorical cohort comparisons.

5. Discussion

This study examined subjective and objective AI literacy among adults sampled from seven districts of Tehran, focusing on

More recent birth year was positively associated with subjective AI literacy, $r_s = .365$, $p < .001$, $n = 395$. In contrast, birth year was not statistically associated with objective AI literacy, $r_s = .020$, $p = .690$, $n = 393$. This pattern was consistent with the categorical cohort analysis and did not depend on representing birth year solely through the prespecified dichotomy.

A separate sensitivity analysis assessed whether exclusion of adult respondents residing outside the seven prespecified sampling districts materially affected the principal subjective–objective association. When these district-protocol deviations were retained in the adult analytic pool, the association remained positive and weak, $r_s = .143$, $p = .003$, $n = 430$, compared with $r_s = .161$, $p = .001$, $n = 389$ in the primary protocol-compliant sample.

The interpretation of a weak positive association between subjective and objective AI literacy was therefore unchanged.

the correspondence between these two forms of assessment and their relationship with birth cohort. Three findings seem important. First, subjective AI-literacy scores were concentrated above the midpoint of the AILS response scale, whereas performance on the objective measure was strongly concentrated at the lower end of its score distribution.

Second, subjective and objective AI literacy were positively but only weakly associated. Third, later-born respondents

reported higher subjective AI literacy, while no statistically detectable cohort difference was observed in objective AI literacy.

Together, these findings demonstrate that conclusions about AI literacy in this sample differed substantially according to the form of assessment used.

The descriptive results revealed different patterns across the two forms of assessment.

The mean overall subjective AI-literacy score was 4.98 on the 1–7 AILS scale, with ethics and evaluation showing the highest mean self-ratings. In contrast, respondents answered an average of 1.31 of the 18 objective AICOS items correctly, and the median objective score was zero.

Approximately 70% of respondents with complete AICOS data obtained no correct answers, while “I do not know” accounted for approximately three quarters of all available item responses. The most notable characteristic of the objective scores was therefore their pronounced concentration at the lower boundary of the instrument.

The distinction between these results is consistent with an important measurement issue in AI-literacy research. The AILS developed by Wang et al. (2023) operationalizes AI literacy through self-reported awareness, use, evaluation, and ethics. Subsequent adaptation research has examined this self-report approach in other linguistic and cultural contexts (Çelebi et al., 2023; Hobeika et al., 2024). Objective approaches address a different measurement target. Chiu et al. (2024), for example, distinguish perceived AI capability assessed through self-report from performance assessed using an objective AI-literacy test. The present findings illustrate the practical importance of preserving this distinction: the characterization of AI literacy in this adult sample differed substantially depending on whether assessment concerned respondents’ perceived capabilities or their performance on an objective knowledge test.

The two forms of assessment were nevertheless not completely unrelated. A statistically significant positive association was observed between subjective and objective AI literacy ($r_s = 0.161$, $p = 0.001$), although its magnitude was small.

Respondents reporting higher subjective AI literacy therefore tended to obtain somewhat higher objective scores, but the monotonic association between the two measures was weak.

This result is consistent with treating subjective and objective assessments as related but non-interchangeable sources of information. It also supports the methodological value of including both forms of assessment when a research question concerns both perceived capability and demonstrated performance.

Importantly, the combination of comparatively higher AILS scores and very low AICOS scores should not be interpreted as direct evidence that respondents systematically overestimated their AI competence.

The two instruments differ in construct coverage, response format, score range, and measurement properties. Their numerical scores therefore cannot be subtracted or directly compared as though they represented corresponding points on a common metric.

Demonstrating overestimation, underestimation, or metacognitive calibration would require a design in which perceived and demonstrated competence are linked at comparable levels—for example, through item-specific confidence judgments paired with corresponding objective responses.

The present data support the narrower conclusion that subjective and objective AI-literacy assessments showed limited correspondence. They do not constitute a direct test of overconfidence, underconfidence, or the Dunning–Kruger effect.

The birth-cohort findings provide a further distinction between subjective and objective assessment. Respondents born in Solar Hijri year 1373 or later reported significantly higher overall subjective AI literacy than earlier-born respondents, with a moderate standardized difference. The same pattern was observed for awareness, use, and evaluation after adjustment for multiple comparisons, whereas the difference in ethics was not statistically significant. The observed cohort difference in subjective

literacy was therefore not uniform across all four AILS dimensions.

This pattern can be considered alongside previous research reporting generation-related differences in engagement with emerging AI technologies. Chan and Lee (2023), for example, reported differences in experiences, perceptions, and intentions regarding generative AI between Gen Z students and Gen X and Millennial teachers. Such evidence provides useful context for the subjective component of the present findings, but attitudes, adoption intentions, technology use, and self-perceived capability are not equivalent to objectively demonstrated AI knowledge. In the present study, later-born respondents reported higher subjective AI literacy, but no statistically detectable difference between the two cohorts was observed in AICOS performance.

The results are also consistent with broader critiques of the assumption that growing up with digital technologies necessarily corresponds to uniformly greater digital competence. Hargittai (2010) documented substantial variation in Internet skills and uses even among young adults commonly characterized as members of the “Net Generation,” while Calderón Gómez (2019) demonstrated the persistence of digital inequalities among young people in relation to differing forms of technological capital. The present findings warrant similar caution in the context of AI literacy. Later-born respondents reported higher perceived capability in several AI-literacy domains, but this pattern was not accompanied by higher performance on the objective measure. Birth cohort should therefore not be treated as a proxy for demonstrated AI knowledge on the basis of these results.

The supplementary continuous birth-year analysis yielded a pattern consistent with the categorical comparison. More recent birth year was associated with higher subjective AI literacy ($r_s = 0.365$, $p < 0.001$), whereas its association with objective AI literacy was close to zero ($r_s = 0.020$, $p = 0.690$). This indicates that the contrasting subjective and objective patterns were also observed when birth year was retained as a continuous

variable rather than represented solely through the prespecified 1373 boundary. However, because the study is cross-sectional, these associations cannot distinguish age, cohort, and period effects. Nor do the present analyses establish why later-born respondents reported higher subjective literacy; possible mechanisms such as differential exposure to AI, frequency of use, educational experience, or confidence were not independently tested.

The extremely low objective scores require particularly cautious interpretation. Although the AICOS correctness score showed high internal consistency in the present sample ($KR-20 = 0.876$), its distribution exhibited a pronounced floor effect. High internal consistency does not establish construct validity, cross-cultural equivalence, or adequate discrimination across the full range of ability. The concentration of scores at or near zero may partly reflect low performance on the knowledge assessed by these items, but it may also have been influenced by item difficulty, linguistic or cultural adaptation, respondents’ familiarity with the content assessed, or the suitability of the short-form instrument for a heterogeneous adult community sample. These possibilities cannot be distinguished with the present data.

Interpretation is further complicated by the addition of an “I do not know” response option to the AICOS items. This option was not part of the original response format and may have influenced response behavior. Although it enabled explicit uncertainty to be distinguished descriptively from selection of an incorrect substantive response, neither response type provides a direct measure of respondents’ metacognitive awareness. The modification also limits direct comparability with administrations using the original AICOS response format.

These measurement considerations limit direct numerical comparisons with AI-literacy findings obtained in other populations. Previous studies have frequently examined educationally selected populations and have used different instruments and scoring procedures. For example, Çelebi et al. (2023) examined a

Turkish adaptation of AILS among adults, Hobeika et al. (2024) examined an Arabic adaptation among university students across four Arab countries, and Chiu et al. (2024) developed an objective AI-literacy assessment for students in Grades 7–9. Such studies provide relevant evidence about approaches to measuring AI literacy across populations and contexts, but their numerical scores are not directly comparable with AICOS scores obtained in the present adult sample. Cross-study interpretation should therefore focus on measurement approaches and patterns rather than treating scores from non-equivalent instruments and populations as a common ranking metric.

More broadly, the findings highlight the importance of specifying what is meant by AI literacy in empirical research. Long and Magerko (2020) conceptualize AI literacy through a set of competencies that enable individuals to critically evaluate AI technologies, communicate and collaborate effectively with AI, and use AI as a tool in different settings. Different instruments operationalize subsets or manifestations of AI literacy in different ways. In the present study, AILS assessed perceived capabilities across awareness, use, evaluation, and ethics, whereas AICOS assessed performance on objective AI-related knowledge items. The weak association between their overall scores indicates that these two operationalizations should not be treated as interchangeable. Researchers should therefore specify whether conclusions concern self-perceived capability, performance on an objective assessment, or both, rather than assuming that different forms of AI-literacy measurement provide equivalent information.

The findings also have implications for educational evaluation, although these implications should remain proportional to the study design. The study did not evaluate an educational intervention and therefore cannot establish which, if any, training approach would improve AI literacy. It does indicate that higher self-reported capability should not automatically be interpreted as higher performance on an objective assessment. Evaluations of future AI-literacy

interventions could therefore benefit from measuring both outcomes and examining whether changes in perceived capability are accompanied by changes in objectively assessed performance.

Several limitations should be considered. First, the sampling design incorporated structured geographic coverage and random selection at the district level. Tehran's 22 municipal districts were grouped into five geographic clusters, from which seven districts were randomly selected according to the prespecified sampling plan. The design also specified equal recruitment from two prespecified birth-cohort strata within each district. However, recruitment of individual respondents within the selected districts was not based on a complete population sampling frame, and the achieved district and cohort allocations did not fully correspond to the planned quotas. Accordingly, although the design incorporated random district selection and structured geographic and cohort coverage, the achieved respondent sample should not be treated as a probability sample of Tehran's adult population. The findings should therefore be generalized to the wider Tehran adult population with caution and should not be interpreted as nationally representative of Iran. Second, the cross-sectional design prevents causal interpretation and does not permit age, cohort, and period effects to be separated.

Third, although the instruments underwent structured translation and cultural-adaptation procedures, full psychometric and cross-cultural validation of the Persian versions was not established. This limitation is particularly consequential for AICOS given the pronounced floor concentration observed in the present sample. Fourth, the restricted objective-score distribution may attenuate or otherwise constrain the observed association between subjective and objective scores.

Consequently, the weak correlation should be interpreted in conjunction with the measurement limitations of the objective score rather than as a precise estimate of the relationship that would necessarily be observed with an instrument providing greater discrimination across the ability range. Fifth, the addition of an "I do not

know” response option modified the original AICOS administration format and may have affected both response selection and comparability with previous administrations.

Future research should replicate the subjective–objective comparison using Persian instruments subjected to more extensive psychometric evaluation and samples representing a broader range of AI-related knowledge and performance. Research specifically concerned with calibration or overestimation should use designs that directly link confidence judgments to corresponding objective items or otherwise establish a defensible common framework for perceived and demonstrated performance. Longitudinal or repeated cross-sectional designs could additionally help distinguish whether birth-related differences in subjective AI literacy reflect cohort-specific experiences, life-course changes, or broader temporal changes in exposure to AI.

Overall, the findings demonstrate that subjective and objective assessments produced substantially different patterns in this sample of adults from seven districts of Tehran. Subjective AI-literacy scores were concentrated above the midpoint of the AILS scale, whereas performance on the administered objective assessment was concentrated at its lower boundary. The two measures showed a weak positive association. Later-born respondents reported higher subjective AI literacy, but neither the categorical cohort comparison nor the continuous birth-year analysis showed evidence of a corresponding advantage in objective performance. The principal contribution of the study is therefore not evidence of an AI-literacy “gap” or systematic overestimation, but evidence that perceived AI-related capability and performance on an objective AI-literacy assessment should be distinguished empirically and interpreted according to the measurement approach used.

5. Conclusion

This study examined subjective and objective AI literacy among adults sampled from seven districts of Tehran, with particular attention to the correspondence

between these two forms of assessment and their relationship with birth cohort. The findings demonstrate the importance of distinguishing between perceived AI-related capability and performance on an objective AI-literacy assessment. Subjective AI-literacy scores were concentrated above the midpoint of the AILS scale, whereas performance on the objective AICOS assessment was strongly concentrated at the lower end of the score distribution. Subjective and objective AI literacy were positively but weakly associated, indicating limited correspondence between the two forms of assessment.

The birth-year analyses revealed different patterns for subjective and objective AI literacy. Later-born respondents reported higher overall subjective AI literacy, with higher scores in awareness, use, and evaluation but no statistically significant difference in ethics. In contrast, no statistically detectable difference in objective AI literacy was observed between the earlier- and later-born cohorts. The supplementary analysis using continuous birth year produced a consistent pattern: more recent birth year was positively associated with subjective AI literacy but was not statistically associated with objective AI literacy. These findings caution against treating more recent birth year or membership in a younger cohort as a proxy for demonstrated AI knowledge.

The study therefore contributes to AI-literacy research by demonstrating the empirical value of distinguishing subjective and objective forms of assessment. Self-report measures provide information about individuals’ perceptions of their AI-related capabilities, whereas objective assessments provide information about performance on the knowledge or competencies they assess. The weak association observed in this study indicates that these forms of evidence should not be treated as interchangeable. Where both perceived capability and demonstrated performance are relevant to the research question, using both forms of assessment may provide a more informative characterization of AI literacy than relying on either alone.

The objective findings nevertheless require particular caution. AICOS scores exhibited a pronounced floor effect, and full psychometric and cross-cultural validation of the Persian short-form instrument was not established. In addition, the researcher-added “I do not know” response option modified the original administration format. Consequently, the observed objective-score distribution cannot be interpreted as a definitive estimate of objective AI literacy among adults in Tehran. Further research is needed to determine the extent to which this distribution reflects performance on the assessed content, item difficulty, characteristics of the Persian adaptation, the modified response format, or other measurement factors.

The scope of the conclusions is also limited by the sampling design. The findings concern 400 adult respondents recruited from seven prespecified districts of Tehran through a geographically structured, quota-based intercept procedure and should not be generalized statistically to the entire adult population of Tehran or Iran. Moreover, the cross-sectional design does not permit causal interpretation of birth-related differences or separation of age, cohort, and period effects. Future research should examine these relationships using broader probability-based samples, Persian-language measures subjected to more extensive psychometric evaluation, and longitudinal or repeated cross-sectional designs. Studies specifically concerned with overestimation or metacognitive calibration should use designs that permit perceived and demonstrated performance to be compared on a defensible common framework rather than inferring calibration from separate instruments.

Within these limitations, the central conclusion is clear: the interpretation of AI literacy depends in part on how it is operationalized and measured. In this sample, higher subjective AI literacy among later-born respondents was not accompanied by a statistically detectable difference in objective AI literacy, and subjective and objective scores showed only a weak positive association. These findings support maintaining a clear conceptual and empirical

distinction between perceived AI-related capability and performance on objective AI-literacy assessments.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

- Alinaghian, A., Safdari Ranjbar, M., & Mohammadi, M. (2023). Designing a policy package for developing artificial intelligence (AI) in Iran. *Iranian Journal of Public Policy*, 9(1), 22–46. <https://doi.org/10.22059/jppolicy.2023.92986>
- Calderón Gómez, D. (2019). Technological capital and digital divide among young people: An intersectional approach. *Journal of Youth Studies*, 22(7), 941–958. <https://doi.org/10.1080/13676261.2018.1559283>
- Carolus, A., Koch, M. J., Straka, S., Latoschik, M. E., & Wienrich, C. (2023). MAILS—Meta AI Literacy Scale: Development and testing of an AI literacy questionnaire based on well-founded competency models and psychological change- and meta-competencies. *Computers in Human Behavior: Artificial Humans*, 1(2), 100014. <https://doi.org/10.1016/j.chbah.2023.100014>
- Çelebi, C., Yılmaz, F., Demir, U., & Karakuş, F. (2023). Artificial intelligence literacy: An adaptation study. *Instructional Technology and Lifelong Learning*, 4(2), 291–306. <https://doi.org/10.52911/ital.1401740>
- Chan, C. K. Y., & Lee, K. K. W. (2023). The AI generation gap: Are Gen Z students more interested in adopting generative AI such as ChatGPT in teaching and learning than their Gen X and millennial generation teachers? *Smart Learning Environments*, 10, 60. <https://doi.org/10.1186/s40561-023-00269-3>
- Chiu, T. K. F., Chen, Y., Yau, K. W., Chai, C.-S., Meng, H., King, I., Wong, S., & Yam, Y. (2024). Developing and validating

- measures for AI literacy tests: From self-reported to objective measures. *Computers and Education: Artificial Intelligence*, 7, 100282.
<https://doi.org/10.1016/j.caeai.2024.100282>
- Grassini, S. (2024). A psychometric validation of the PAILQ-6: Perceived Artificial Intelligence Literacy Questionnaire. In *Proceedings of the 13th Nordic Conference on Human-Computer Interaction*. Association for Computing Machinery.
<https://doi.org/10.1145/3679318.3685359>
- Hargittai, E. (2010). Digital na(t)ives? Variation in Internet skills and uses among members of the “Net Generation.” *Sociological Inquiry*, 80(1), 92–113.
<https://doi.org/10.1111/j.1475-682X.2009.00317.x>
- Hobeika, E., Hallit, R., Malaeb, D., Sakr, F., Dabbous, M., Merdad, N., Rashid, T., Amin, R., Jebreen, K., Zarrouq, B., Alhuwailah, A., Shuwiekh, H. A. M., Hallit, S., Obeid, S., & Fekih-Romdhane, F. (2024). Multinational validation of the Arabic version of the Artificial Intelligence Literacy Scale (AILS) in university students. *Cogent Psychology*, 11(1), 2395637.
<https://doi.org/10.1080/23311908.2024.2395637>
- Hornberger, M., Bewersdorff, A., & Nerdel, C. (2023). What do university students know about artificial intelligence? Development and validation of an AI literacy test. *Computers and Education: Artificial Intelligence*, 5, 100165.
<https://doi.org/10.1016/j.caeai.2023.100165>
- Jin, Y., Martinez-Maldonado, R., Gašević, D., & Yan, L. (2025). GLAT: The generative AI literacy assessment test. *Computers and Education: Artificial Intelligence*, 9, 100436.
<https://doi.org/10.1016/j.caeai.2025.100436>
- Laupichler, M. C., Aster, A., Haverkamp, N., & Raupach, T. (2023). Development of the “Scale for the Assessment of Non-Experts’ AI Literacy”—An exploratory factor analysis. *Computers in Human Behavior Reports*, 12, 100338.
<https://doi.org/10.1016/j.chbr.2023.100338>
- Long, D., & Magerko, B. (2020). What is AI literacy? Competencies and design considerations. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (pp. 1–16). Association for Computing Machinery.
<https://doi.org/10.1145/3313831.3376727>
- Markus, A., Carolus, A., & Wienrich, C. (2025). Objective measurement of AI literacy: Development and validation of the AI Competency Objective Scale (AICOS). *Computers and Education: Artificial Intelligence*, 9, 100485.
<https://doi.org/10.1016/j.caeai.2025.100485>
- Ng, D. T. K., Wu, W., Leung, J. K. L., Chiu, T. K. F., & Chu, S. K. W. (2024). Design and validation of the AI literacy questionnaire: The affective, behavioural, cognitive and ethical approach. *British Journal of Educational Technology*, 55(3), 1082–1104. <https://doi.org/10.1111/bjet.13411>
- Van Deursen, A. J. A. M., & Van Dijk, J. A. G. M. (2011). Internet skills and the digital divide. *New Media & Society*, 13(6), 893–911.
<https://doi.org/10.1177/1461444810386774>
- Van Dijk, J. A. G. M. (2012). The evolution of the digital divide: The digital divide turns to inequality of skills and usage. In J. Bus, M. Crompton, M. Hildebrandt, & G. Metakides (Eds.), *Digital Enlightenment Yearbook 2012* (pp. 57–78). IOS Press.
- Van Dijk, J. A. G. M. (2017). Digital divide: Impact of access. In P. Rössler, C. A. Hoffner, & L. van Zoonen (Eds.), *The International Encyclopedia of Media Effects*. Wiley.
<https://doi.org/10.1002/9781118783764.wbieme0043>
- Wang, B., Rau, P.-L. P., & Yuan, T. (2023). Measuring user competence in using artificial intelligence: Validity and reliability of artificial intelligence literacy scale. *Behaviour & Information Technology*, 42(9), 1324–1337.
<https://doi.org/10.1080/0144929X.2022.2072768>
- Zhang, S., Xiao, R., Botelho, A. F., Liao, G., Chiu, T. K. F., Stamper, J., & Koedinger, K. R. (2026). How to assess AI literacy: Misalignment between self-reported and objective-based measures. In *Proceedings of the LAK26: 16th International Learning Analytics and Knowledge Conference* (pp. 405–414). Association for Computing Machinery.
<https://doi.org/10.1145/3785022.3785088>