From Knowledge to Intention: Developing and Validating a Scale for K-12 Teachers' Readiness to Teach AI

Hasan Tokatlı

htokatli@yildiz.edu.tr

Department of Informatics, Yıldız Technical University, Istanbul, Türkiye https://orcid.org/0000-0002-9231-4191

M. Fatih Erkoç

Department of Computer Education and Instructional Technology, Faculty of Education, Yıldız Technical University, Istanbul, Türkiye https://orcid.org/0000-0002-8278-2805

Research Article

Keywords: readiness to teach AI, behavioral intention, scale development, TPACK, Theory of Planned Behavior, K-12 education

Posted Date: April 2nd, 2026

DOI: https://doi.org/10.21203/rs.3.rs-9291938/v1

Additional Declarations: The authors declare no competing interests.

Abstract

The growing integration of artificial intelligence (AI) into society has increased demand for K-12 AI education, placing information technology (IT) teachers at the forefront of curriculum delivery. However, validated instruments assessing IT teachers' readiness and motivation toward teaching AI remain scarce. This study developed and validated the Readiness to Teach AI and Behavioral Intention Scale (RTAI-BIS), grounded in the Technological Pedagogical Content Knowledge (TPACK) framework and the Theory of Planned Behavior (TPB). An initial pool of 294 items was refined through literature review, expert panel evaluation (n = 6), and pilot testing, yielding a 45-item candidate scale. Exploratory Factor Analysis (EFA) with 392 IT teachers in Turkish schools yielded a five-factor structure explaining 73.79% of variance. Confirmatory Factor Analysis (CFA) with an independent sample (n = 258) validated a refined four-factor, 26-item model with acceptable fit (Comparative Fit Index [CFI] = .943, root mean square error of approximation [RMSEA] = .079, standardized root mean square residual [SRMR] = .061). The final four factors are Technological, Pedagogical, and Content Knowledge for Teaching AI (TPACK-TAI), Attitude Toward AI (ATA), Technological Knowledge for Teaching AI (TK-TAI), and Disposition Toward Teaching AI (DTAI), with two knowledge factors capturing related but distinguishable aspects of instructional readiness. Reliability analyses demonstrated strong internal consistency (Cronbach's α = .921–.974). The RTAI-BIS offers a psychometrically sound tool for assessing IT teachers' readiness to teach AI and their behavioral intentions, with applications in teacher training and curriculum implementation.

1. Introduction

Artificial intelligence (AI) is reshaping work, knowledge production, and the skill expectations placed on education systems (Di Battista et al., 2023; Frey & Osborne, 2017). Schools feel this shift acutely. K-12 systems face growing pressure to prepare students to understand AI's logic, applications, and social implications (Bessen, 2018; Martinez, 2018; Panth & Maclean, 2020).

1.1 AI Education at the K-12 Level

Since 2019, K-12 AI education has shifted from a peripheral topic to an emerging curriculum field, supported by frameworks that specify what students should know about AI and how those ideas can be taught in age-appropriate ways (Kim et al., 2021b; Long & Magerko, 2020; Touretzky et al., 2019). Reviews of this literature converge on three themes. First, K-12 AI education is expected to address foundational concepts, machine learning, and the social consequences of AI rather than narrow technical training alone (Casal-Otero et al., 2023; Rizvi et al., 2023; Su et al., 2024). Second, instruction tends to rely on applied, student-centered approaches such as project-based learning and collaborative design tasks (Ng et al., 2023; Sanusi et al., 2023; Su et al., 2022). Third, and most critically for implementation, teacher preparedness remains a persistent bottleneck: teachers need domain knowledge, pedagogical strategies, and confidence to teach AI, yet these capacities are unevenly developed and rarely assessed through robust instruments (Kim et al., 2021a; Ng et al., 2023; Yue et al., 2024). These global patterns are evident in Turkey, where the broader policy movement identified by

UNESCO (2021) as dependent on both curriculum design and teacher capacity is already underway. The National Artificial Intelligence Strategy 2021–2025 (T.C. Cumhurbaşkanlığı Dijital Dönüşüm Ofisi [CBDDO], 2021) and the AI Applications curriculum issued by the Ministry of National Education (T.C. Millî Eğitim Bakanlığı [MEB], 2023) indicate that AI teaching is no longer a speculative future concern, but an implementation challenge for schools and teachers.

1.2 The Critical Role of Teachers

Teachers are central to whether AI curricula become meaningful classroom practice. In many school systems, and particularly in the Turkish context, information technology (IT) teachers are among the most likely implementers of formal AI instruction. Their preparedness therefore has both a competence dimension, concerning whether they believe they have the knowledge and pedagogical capacity to teach AI, and an intention dimension, concerning whether they are willing and motivated to do so.

Existing studies point to this dual challenge, but they do so unevenly. Research has examined teachers' motivation and behavioral intention toward AI teaching (Chai et al., 2024; Ayanwale et al., 2022; Sanusi et al., 2024), their AI-related knowledge or TPACK readiness (Kim et al., 2021a; Yue et al., 2024), and instruments for adjacent constructs such as using AI as a pedagogical tool, general AI self-efficacy, or teacher acceptance of AI (Celik, 2023; Guo et al., 2025; Ramazanoğlu & Akın, 2025; Wang & Chuang, 2024). However, these strands remain largely separate, and none addresses readiness to teach AI as curriculum content with an integrated measurement approach. Cross-national evidence reinforces this gap: Du et al. (2023) found that even teachers with moderate AI literacy felt unprepared to adapt AI curricula to their classrooms, and Addo (2023) reported similar barriers among UK teachers — highlighting the disconnect between general awareness and teaching-specific readiness. To our knowledge, no validated instrument jointly measures knowledge-based readiness and behavioral intention for teaching AI as a school subject — a gap especially acute in Turkey, where policy attention to AI education has advanced more quickly than validated assessment tools.

1.3 Purpose of the Study

Accordingly, this study aimed to develop and validate a psychometrically sound instrument, the Readiness to Teach AI and Behavioral Intention Scale (RTAI-BIS), for measuring IT teachers' readiness to teach AI and their behavioral intentions toward doing so.

The study was guided by the following research questions:

1. What is the factor structure of the RTAI-BIS as determined by exploratory factor analysis? 2. To what extent does confirmatory factor analysis support the factor structure identified through EFA? 3. Does the RTAI-BIS demonstrate acceptable levels of reliability?

The RTAI-BIS is designed to support readiness profiling of teacher populations, evaluation of professional development interventions, and cross-regional comparison of teacher preparedness for AI instruction.

2. Theoretical Framework

The RTAI-BIS was informed by two complementary frameworks: the Technological Pedagogical Content Knowledge (TPACK) framework and the Theory of Planned Behavior (TPB). Together, they provide a basis for conceptualizing teacher preparedness as both a competence issue and an intention issue.

2.1 Technological Pedagogical Content Knowledge (TPACK)

The TPACK framework (see Fig. 1) extends pedagogical content knowledge by emphasizing that effective teaching with technology requires the coordinated use of content knowledge, pedagogical knowledge, and technological knowledge (Koehler & Mishra, 2009; Mishra & Koehler, 2006; Shulman, 1986, 1987). In the context of AI education, this means that teachers need not only to understand AI concepts, but also to select suitable pedagogical approaches and use technological tools in ways that make those concepts teachable to K-12 learners.

TPACK has become a common lens for examining teacher readiness in technology-rich settings (Koehler & Mishra, 2009; Kim et al., 2021a). In AI education specifically, Kim et al. (2021b) argued that TPACK can frame the knowledge teachers need for K-12 AI instruction, and Yue et al. (2024) showed that teachers' AI-related content and technological knowledge may lag behind their general pedagogical confidence. At the same time, newer AI-focused extensions such as Intelligent-TPACK and AI-TPACK often target teachers' capacity to use AI as a teaching tool rather than to teach AI as curricular content (Celik, 2023). The RTAI-BIS adopts TPACK in this latter, subject-teaching sense: the readiness component of the RTAI-BIS was designed to capture teachers' perceived readiness to teach AI concepts, practices, and applications.

Within the TPACK model, technological knowledge (TK) is one component of the broader TPACK framework rather than a stand-alone representation of integrated instructional knowledge (Mishra & Koehler, 2006). However, in emerging subject areas such as AI education — where the tools themselves are the curricular content — TK-related competencies may be empirically separable from the broader pedagogical-content integration that TPACK represents. Whether this separation manifests in the present context is treated as an empirical question.

2.2 Theory of Planned Behavior (TPB)

The TPB explains behavior through the mediating role of intention, which is shaped by attitude toward the behavior, subjective norms, and perceived behavioral control (Ajzen, 1991, 2020). Applied to AI teaching, these components concern whether teachers see teaching AI as worthwhile, perceive social support or pressure around it, and believe they have sufficient capability and resources to do it.

This framework is relevant because intention-based models consistently explain teachers' adoption of new technologies and practices (Ajzen, 2011, 2020; Davis, 1989; Scherer et al., 2019; Venkatesh et al., 2003). TPB has been applied to domain-specific teaching intentions — for example, Lin and Williams (2016) modeled preservice teachers' STEM teaching intentions and found perceived behavioral control and subjective norms to be the strongest predictors. In AI education specifically, prior work has linked intention to attitudes, efficacy beliefs, confidence, and perceived relevance (Ayanwale et al., 2022; Chai et al., 2024; Sanusi et al., 2024). Accordingly, the behavioral-intention component of the RTAI-BIS was designed with TPB as its conceptual basis.

2.3 Integrating TPACK and TPB: The Conceptual Foundation of the RTAI-BIS

TPACK and TPB were used together because readiness to teach AI cannot be understood only as knowledge possession or only as willingness to act. TPACK helps specify what teachers need to know to teach AI, whereas TPB helps explain whether they are disposed to do so within their institutional context. A similar dual-framework strategy has been employed by Habibi et al. (2023), who integrated TPB with TPACK to model preservice teachers' technology integration intentions and found that the combination explained substantially more variance than either framework alone. The initial item pool was therefore constructed to represent both readiness-related and TPB-informed intention-related dimensions (see Fig. 2 for the theoretical-to-empirical mapping). How the TPB-informed items performed during scale refinement is examined in the Discussion (Section 5.1).

3. Methodology

3.1 Research Design

This study employed a scale development research design to create a valid and reliable instrument for measuring IT teachers' readiness to teach AI and their behavioral intentions. The scale development process followed established guidelines in the literature (Büyüköztürk, 2002; DeVellis, 2017; Güngör, 2016; Koyuncu & Kılıç, 2019) and comprised three major phases: item pool development, validity studies, and reliability studies. The overall scale development procedure is illustrated in Fig. 3.

3.2 Item Pool Development

3.2.1 Literature Review

In the initial phase, national and international AI education curricula (both formal and informal), AI textbooks, and scale items used in related empirical studies were systematically reviewed. This review served to identify the content domain and inform the generation of candidate items aligned with the TPACK and TPB frameworks.

3.2.2 Item Generation

Based on the literature review, a preliminary pool of 294 items was generated. The researchers refined this pool by eliminating overlapping or redundant items while ensuring alignment with the theoretical foundations, yielding a draft pool of 52 items for expert review.

3.2.3 Expert Review

The 52-item draft was submitted to a panel of six academic experts for content validity evaluation, linguistic appropriateness assessment, and face validity review. Expert panel details are presented in Table 1.

Table 1

Expert Panel Composition

Expert	Faculty / Department	Title
1	Faculty of Education, Computer Education and Instructional Technology	Professor
2	Faculty of Education, Computer Education and Instructional Technology	Associate Professor
3	Faculty of Education, Computer Education and Instructional Technology	Research Assistant (PhD)
4	Faculty of Education, Educational Sciences	Professor
5	Faculty of Education, Educational Sciences	Associate Professor
6	Faculty of Education, Foreign Languages Education (Language Specialist)	Assistant Professor

Following expert review, seven items were removed and modifications were applied to selected items based on expert recommendations, resulting in a 45-item candidate scale.

3.2.4 Pilot Testing

A pilot study was conducted to further assess face validity. The 45-item candidate scale was converted into a Google Forms questionnaire and administered online to 40 senior preservice IT teachers enrolled at Yıldız Technical University, Department of Computer Education and Instructional Technology. Of the 40 participants, 10 provided detailed and usable written feedback on item clarity and relevance. Based on this feedback, final adjustments were made to the candidate scale, which retained its 45-item structure. Although the pilot sample comprised preservice teachers rather than practicing IT teachers, the pilot's purpose was limited to face validity and item clarity assessment, for which this population was considered appropriate.

3.3 Participants

Two independent samples of IT teachers serving in Turkish public and private schools were recruited using non-probability snowball sampling (Table 2). Because no centralized registry of IT teachers exists in Turkey, snowball sampling was selected to maximize geographic reach within this dispersed population. Participants responded to the online questionnaire distributed via electronic channels. Across both study phases combined, respondents represented 77 of Turkey's 81 provinces (71 in the EFA sample, 58 in the CFA sample), indicating that the sampling chains extended well beyond localized networks.

Study 1 (Exploratory Factor Analysis)

A total of 392 IT teachers participated in the EFA phase.

Study 2 (Confirmatory Factor Analysis)

A total of 258 IT teachers participated in the CFA phase. This sample exceeds the minimum of 200 cases recommended for ML-based CFA (Kline, 2023) and provides approximately four observations per freely estimated parameter in the final model (Hair et al., 2019).

Table 2

Participant Demographics by Study Phase

Variable	Category	Study 1 — EFA (n = 392)		Study 2 — CFA (n = 258)	
		n	%	n	%
Gender	Female	194	49.49	134	51.94
Gender	Male	198	50.51	124	48.06
School Type	Public School	257	65.56	183	70.93
School Type	Private School	135	34.44	75	29.07
Years of Experience	0–5 years	151	38.52	91	35.27
Years of Experience	6–10 years	105	26.79	61	23.64
Years of Experience	11–15 years	57	14.54	45	17.44
Years of Experience	16–20 years	56	14.29	47	18.22
Years of Experience	21 + years	23	5.87	14	5.43
Prior AI Training	Yes	199	50.77	134	51.94
Prior AI Training	No	193	49.23	124	48.06

Note. The two samples were independently recruited via snowball sampling and did not overlap. The demographic distributions were comparable across the two phases, supporting the suitability of independent-sample cross-validation.

3.4 Data Collection Instrument

The candidate scale comprised 45 items. Most items were rated on a 5-point agreement scale (1 = Strongly Disagree, 2 = Disagree, 3 = Undecided, 4 = Agree, 5 = Strongly Agree). However, the five AI Interest and Engagement (AIE) items (M41–M45) used a 5-point frequency format (1 = Never, 2 = Rarely, 3 = Sometimes, 4 = Often, 5 = Very Often), because they were intended to capture behavioral engagement with AI rather than attitudinal endorsement. This mixed response-format design was retained in both study administrations and is considered in interpreting the CFA results (see Section 4.2).

Items were designed to assess two overarching dimensions: (a) readiness to teach AI, informed by the TPACK framework, and (b) behavioral intentions related to teaching AI, informed by the TPB. Within the TPB-informed dimension, the initial item pool included items operationalizing all three TPB components. These were attitude toward the behavior (e.g., beliefs about the value of AI education for students), subjective norms (e.g., perceived expectations of colleagues and administrators regarding teaching AI), and perceived behavioral control (e.g., self-efficacy beliefs and perceived resource access for AI instruction). The instrument also included a personal information form collecting demographic data.

The scale items were originally developed in Turkish. For the purposes of international dissemination, a forward-backward translation procedure aligned with the ITC Guidelines for Translating and Adapting Tests (International Test Commission, 2018) was conducted to produce an English version of the scale.

3.5 Data Analysis

The study used IBM SPSS v25 for exploratory factor analysis and IBM SPSS AMOS v26 for confirmatory factor analysis.

Common Method Bias. Because the study relied on a single self-report instrument administered at one time point, common method bias (CMB) is a potential concern (Podsakoff et al., 2003). Rather than relying on Harman's single-factor test, which has been shown to be an insensitive diagnostic (Podsakoff et al., 2003), CMB risk was evaluated through converging structural evidence. Specifically, the multi-factor CFA solution, HTMT values below .90, and AVE values above recommended thresholds were collectively examined to assess whether a single-method factor could plausibly account for the observed correlations.

Exploratory Factor Analysis (EFA). Prior to factor extraction, the analysis examined item-total correlations to identify items with insufficient discrimination. The Kaiser-Meyer-Olkin (KMO) measure and Bartlett's Test of Sphericity assessed sampling adequacy. The study employed principal component analysis for factor extraction. Principal component analysis was selected for data reduction at the exploratory stage. With communalities consistently above .50 and the majority above .70, PCA and common factor methods such as principal axis factoring are expected to yield convergent solutions (Costello & Osborne, 2005; Fabrigar et al., 1999). The number of factors was determined based on the Kaiser criterion (eigenvalues > 1) and scree plot analysis. Direct Oblimin rotation was applied given the theoretical expectation of correlated factors. Items were retained based on factor loading thresholds (≥ .40) and the absence of cross-loadings exceeding .10 difference between factors (Koyuncu & Kılıç, 2019; Yaşlıoğlu, 2017). To verify that the factor structure was not contingent on the extraction method, the analysis was replicated using Principal Axis Factoring (PAF); results are reported in Section 4.1.5.

Confirmatory Factor Analysis (CFA). The factor structure identified through EFA was tested using CFA with the independent sample. CFA was conducted using Maximum Likelihood (ML) estimation. Prior to analysis, univariate and multivariate normality were examined. Univariate skewness values ranged from – 1.97 to – 0.42 and kurtosis values from – 0.75 to 3.23, all within the thresholds recommended by Kline (2023; |skewness| < 2, |kurtosis| < 7). Although Mardia's multivariate kurtosis was elevated (normalized CR = 55.11), simulation research has demonstrated that ML estimation yields robust parameter estimates with 5-category ordinal indicators and sample sizes exceeding 200 (Rhemtulla et al., 2012; Li, 2016; Beauducel & Herzberg, 2006). To account for potential chi-square inflation, Bollen-Stine bootstrap (500 resamples) was applied (Bollen & Stine, 1992). As a further robustness check, the CFA model was re-estimated on a polychoric covariance matrix (Tucker's congruence > .99, confirming ML robustness to ordinal measurement).

The study evaluated model fit using multiple indices: RMSEA, SRMR, GFI, NFI, TLI, CFI, IFI, and χ^2/df (Hu & Bentler, 1999; Kline, 2023). A .70 threshold for standardized regression weights (SRW) was adopted as the primary retention criterion, corresponding to approximately 50% shared variance between an item and its factor (Hair et al., 2019). This threshold is also consistent with the AVE ≥ .50 convergent-validity requirement (Fornell & Larcker, 1981). Re-specification followed a sequential, one-item-at-a-time protocol: after each model estimation, the item with the lowest SRW was identified and evaluated for removal (Koyuncu & Kılıç, 2019; Kline, 2023). Removal decisions were guided primarily by SRW magnitude but also informed by construct alignment and parsimony; items that weakened the definitional focus of their target factor were removed even when their loading was near the threshold. When successive removal left a factor without a stable set of adequately performing indicators, that factor was not retained in the final model. Convergent validity was assessed through Average Variance Extracted (AVE) values, with .50 as the minimum threshold (Fornell & Larcker, 1981; Yaşlıoğlu, 2017). Discriminant validity was evaluated using the Heterotrait-Monotrait (HTMT) ratio of correlations, with values below .90 indicating adequate discriminant validity (Henseler et al., 2015). Configural invariance was examined by fitting the CFA model separately to gender and school type subgroups.

Reliability Analysis. The study assessed internal consistency using Cronbach's alpha (α) for both the overall scale and individual subscales, with .70 as the minimum acceptable threshold (Güngör, 2016; Nunnally & Bernstein, 1994). Composite Reliability (CR) was computed as $CR = (\Sigma\lambda)^2 / [(\Sigma\lambda)^2 + \Sigma(1 - \lambda^2)]$, where λ denotes standardized factor loadings (Fornell & Larcker, 1981), with .70 as the minimum threshold (Yaşlıoğlu, 2017). McDonald's omega (ω) was additionally computed as a model-based reliability coefficient that does not assume tau-equivalence (Revelle & Zinbarg, 2009). Item-total correlations and inter-subscale correlations were also examined.

3.6 Ethical Considerations

This study received ethical approval from the Yıldız Technical University Social and Human Sciences Research Ethics Committee (Report No: 20240402851, Decision No: 2024.04; date: April 1, 2024). All participants were informed about the purpose of the study, the voluntary nature of participation, and the confidentiality of their responses. Informed consent was obtained electronically prior to questionnaire completion. No personally identifiable information was collected.

3.7 Data Availability

The anonymized dataset generated during this study is available from the corresponding author upon reasonable request.

4. Results

4.1 Exploratory Factor Analysis

4.1.1 Preliminary Analyses

Item-total correlations were computed to assess each item's consistency with the overall scale. Four items (M7, M31, M32, M5) demonstrated very low correlation values and were removed from further analyses, leaving 41 items. The Cronbach's alpha coefficient for the remaining 41 items was .976 (alpha was not computed at the initial 45-item stage, as it would have been deflated by items with near-zero correlations).

Common Method Bias Assessment. Because all data were collected through a single self-report instrument, CMB was evaluated through converging structural evidence (Podsakoff et al., 2003): (a) the EFA distributed variance across five factors rather than one, (b) all HTMT values remained below .90 (see Section 4.2), and (c) AVE values exceeded .50 for all factors. These results suggest that common method variance does not provide a plausible single-factor explanation for the observed correlations; procedural remedies for future administrations are discussed in Section 5.5.

The Kaiser-Meyer-Olkin (KMO) test and Bartlett's Test of Sphericity were conducted to evaluate the suitability of the data for factor analysis (Table 3).

Table 3

KMO and Bartlett's Test of Sphericity Results

Test	Value
KMO Measure of Sampling Adequacy	.966
Bartlett's Test of Sphericity — Approx. χ²	17437.946
Bartlett's Test of Sphericity — df	820
Bartlett's Test of Sphericity — p	< .001

Note. N = 392.

Both indices confirmed that the data were suitable for factor analysis. The KMO value (.966) well exceeded the recommended .60 threshold, and Bartlett's test was statistically significant (Büyüköztürk, 2002; Koyuncu & Kılıç, 2019).

4.1.2 Factor Extraction

Principal Component Analysis was conducted with the remaining 41 items. The Kaiser criterion identified five factors with eigenvalues greater than 1.0 (Table 4). The scree plot analysis corroborated this five-factor solution.

Table 4

Eigenvalues and Explained Variance for the First 10 Components

Component	Eigenvalue	% of Variance	Cumulative %
1	21.274	51.887	51.887
2	4.058	9.896	61.784
3	1.996	4.869	66.653
4	1.711	4.173	70.827
5	1.216	2.966	73.793
6	0.948	2.313	76.105
7	0.752	1.835	77.940
8	0.674	1.645	79.584
9	0.625	1.524	81.109
10	0.573	1.399	82.507

The scree plot (Fig. 4) visually confirmed the five-factor solution, with a clear inflection point after the fifth component.

Communality values for all 41 items were above .50, indicating that the extracted factors adequately accounted for the variance in each item (Erkuş, 2019; Yaşlıoğlu, 2017).

4.1.3 Factor Rotation and Item Retention

The unrotated component matrix was examined, and four items (M6, M8, M13, M15) that did not meet the retention criteria (factor loading ≥ .40 on any factor, or cross-loading difference < .10) were removed.

Direct Oblimin oblique rotation was applied to the remaining 37 items, based on the theoretical expectation of inter-factor correlations. The rotated component matrix is presented in Table 5.

Table 5

Rotated Component Matrix (Direct Oblimin) — Factor Loadings

Item	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5
M28	.889				
M23	.882				
M26	.851				
M27	.848				
M24	.831				
M29	.823				
M30	.800				
M25	.789				
M22	.785				
M18	.762				
M17	.759				
M20	.708				
M19	.670				
M16	.649				
M36		.925			
M37		.898			
M35		.870			
M34		.854			
M38		.845			
M39		.830			
M33		.822			
M40		.785			
M43			.839		
M42			.820		
M44			.763		
M45			.733		
M41			.719		

Item	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5
M3				.736	
M2				.734	
M4				.726	
M1				.699	
M9				.469	
M14					−.709
M10					−.681
M12					−.664
M21					−.662
M11					−.478

Note. Only primary factor loadings are displayed. Factor loadings < .40 are suppressed for clarity. Negative loadings on Factor 5 reflect the oblique rotation direction and do not indicate reverse-scored items.

Table 6

Inter-Factor Correlation Matrix

F1	F1	F2	F3	F4	F5
	1.000				
F2	.517	1.000			
F3	.506	.315	1.000		
F4	.377	.475	.273	1.000	
F5	−.555	−.317	−.299	−.202	1.000

The inter-factor correlations (Table 6) confirmed the appropriateness of the oblique rotation method.

4.1.4 EFA Summary

The EFA yielded a five-factor structure with 37 items, explaining 73.79% of the total variance (Table 7). Factor labels were assigned based on the theoretical relationships of the items within each factor. This solution was treated as a candidate structure; full reliability analyses were conducted on the CFA-confirmed model (see Section 4.3).

Table 7

EFA Factor Structure Summary

Factor	Label	No. of Items	Item Codes	% Variance
1	Technological, Pedagogical, and Content Knowledge for Teaching AI (TPACK-TAI)	14	M16, M17, M18, M19, M20, M22, M23, M24, M25, M26, M27, M28, M29, M30	51.887
2	Attitude Toward AI (ATA)	8	M33, M34, M35, M36, M37, M38, M39, M40	9.896
3	AI Interest and Engagement (AIE)	5	M41, M42, M43, M44, M45	4.869
4	Disposition Toward Teaching AI (DTAI)	5	M1, M2, M3, M4, M9	4.173
5	Technological Knowledge for Teaching AI (TK-TAI)	5	M10, M11, M12, M14, M21	2.966
Total		37		73.793

4.1.5 Extraction Method Robustness Check

To confirm that the factor structure was not an artifact of the extraction method, the EFA was replicated using Principal Axis Factoring (PAF) with Direct Oblimin rotation. The PAF solution recovered the same five-factor structure with identical item-to-factor assignments. Tucker's congruence coefficients between the PCA and PAF pattern matrices ranged from .92 to .99, exceeding the .85 threshold for fair similarity (Lorenzo-Seva & ten Berge, 2006). The PAF solution explained 61.07% of total variance, consistent with the expected reduction when modeling common variance only (Costello & Osborne, 2005; Fabrigar et al., 1999).

4.2 Confirmatory Factor Analysis

The five-factor, 37-item structure from EFA was tested with the independent CFA sample ($n = 258$) using IBM SPSS AMOS v26. Item retention followed the sequential re-specification procedure described in Section 3.5. At each step, the item with the lowest SRW was identified, evaluated for both statistical performance and substantive alignment with the target construct, and removed before re-estimating the model. This process was repeated until all remaining items met the .70 threshold and model fit indices reached acceptable levels. Across successive re-specifications, a total of 11 items were removed in the following order: M4, M44, M45, M41, M42, M43, M39, M9, M34, M16, and M19. Table 8 summarizes the removed items, their original EFA factor assignments, and the removal rationale.

Table 8

Items Removed During CFA Re-Specification

Order	Item	Original EFA Factor	SRW at Removal	Primary Removal Basis
1	M4	DTAI (F4)	.531	Low loading; less aligned with retained DTAI core
2	M44	AIE (F3)	.594	Low loading; personal AI-use behavior; frequency format
3	M45	AIE (F3)	.584	Low loading; personal AI-use behavior; frequency format
4	M41	AIE (F3)	.625	Low loading; personal AI-use behavior; frequency format
5	M42	AIE (F3)	.963	Personal AI-use behavior; frequency format; factor collapsing
6	M43	AIE (F3)	.924	Final AIE indicator; factor eliminated
7	M39	ATA (F2)	.919	Content overlap with retained ATA items
8	M9	DTAI (F4)	.767	Conditional willingness; less aligned with retained DTAI core
9	M34	ATA (F2)	.916	Peripheral to retained ATA definition
10	M16	TPACK-TAI (F1)	.722	Broader AI knowledge; less teaching-specific
11	M19	TPACK-TAI (F1)	.734	Broader societal awareness; less teaching-specific

Note. SRW = standardized regression weight at the model estimation step immediately preceding the item's removal. Items 1–4 fell below the .70 threshold. Items 5–11 were removed primarily on conceptual grounds, as detailed below.

The removals followed three patterns: (a) all five AIE items (M41–M45) measured personal AI tool-use frequency rather than teaching-specific capability, and their frequency response format introduced heterogeneity (Podsakoff et al., 2003), leading to elimination of the AIE factor; (b) M4 and M9 originated from TPB components whose content diverged from the retained DTAI core; and (c) the remaining items (M34, M39, M16, M19) reflected diffuse evaluations or broad knowledge peripheral to the target constructs. Theoretical implications of these patterns are discussed in Section 5.1.

The remaining four factors retained their conceptual coherence and all item SRWs in the final 26-item model exceeded .70 (see Table 13). Because the AIE factor was eliminated, the surviving factors were renumbered: TK-TAI moved from EFA Factor 5 to CFA Factor 3, and DTAI moved from EFA Factor 4 to CFA Factor 4.

Model fit was further improved by examining modification indices. The largest modification index indicated a covariance between the error terms of items M20 ("I can explain how AI applications work using examples") and M22 ("I can explain the strengths and weaknesses of AI technology"), both of which loaded on the TPACK-TAI factor and shared overlapping content related to AI content knowledge. This single error covariance was specified based on both statistical and substantive grounds. No further modifications were warranted, as remaining modification indices did not indicate theoretically justifiable re-specifications. The final model ($\chi^2 = 759.910$, df = 292, $p < .001$) is presented in Fig. 5, and model fit indices are summarized in Table 9.

Table 9
CFA Model Fit Indices

Index	Value	Acceptable Threshold	Reference
$\chi^2(292)$	759.910, $p < .001$	—	—
χ^2/df	2.602	< 5	Güngör (2016)
RMSEA	.079, 90% CI [.072, .086]	≤ .08	Koyuncu & Kılıç (2019)
SRMR	.061	≤ .08	Yaşlıoğlu (2017)
GFI	.810	≥ .90	Güngör (2016)
NFI	.911	≥ .90	Koyuncu & Kılıç (2019)
TLI	.937	≥ .90	Yaşlıoğlu (2017)
CFI	.943	≥ .90	Güngör (2016)
IFI	.943	≥ .90	Koyuncu & Kılıç (2019)

All primary model fit indices (CFI, TLI, RMSEA, SRMR) were within acceptable ranges, supporting the construct validity of the four-factor, 26-item structure. The RMSEA point estimate (.079) fell within the adequate-fit range, and its 90% confidence interval [.072, .086] indicated that the upper bound remained below .10, reinforcing the acceptability of the model fit. GFI (.810) fell below .90 but is no longer recommended as a primary fit index (Hu & Bentler, 1999; Sharma et al., 2005); model adequacy rests on the primary indices above. Bollen-Stine bootstrap (500 resamples) similarly yielded $p < .001$, consistent with the known sensitivity of χ^2 to sample size.

Table 10

Average Variance Extracted (AVE) Values

Factor	AVE
Factor 1 — TPACK-TAI	.756
Factor 2 — ATA	.831
Factor 3 — TK-TAI	.700
Factor 4 — DTAI	.840
Scale Average	.782

All AVE values exceeded the .50 threshold, indicating adequate convergent validity (Yaşlıoğlu, 2017).

Discriminant validity was further evaluated using the Heterotrait-Monotrait (HTMT) ratio of correlations (Henseler et al., 2015). HTMT values are presented in Table 11.

Table 11

HTMT Discriminant Validity Matrix

TPACK-TAI	TPACK-TAI	ATA	TK-TAI	DTAI
	—			
ATA	.590	—		
TK-TAI	.864	.537	—	
DTAI	.503	.799	.471	—

All HTMT values were below the conservative .90 threshold, supporting discriminant validity across all factor pairs. The highest HTMT value was observed between TPACK-TAI and TK-TAI (.864), which is theoretically expected given that both factors capture knowledge-based dimensions of the TPACK framework. All remaining HTMT values were below the stricter .85 criterion.

4.2.1 CFA Summary

The CFA confirmed a four-factor structure with 26 items. The final factor structure is presented in Table 12.

Table 12

Final Scale Structure After CFA

Factor	Label	No. of Items	Item Codes	Std. Regression Weights
1	Technological, Pedagogical, and Content Knowledge for Teaching AI (TPACK-TAI)	12	M17, M18, M20, M22, M23, M24, M25, M26, M27, M28, M29, M30	.705–.945
2	Attitude Toward AI (ATA)	6	M33, M35, M36, M37, M38, M40	.897–.927
3	Technological Knowledge for Teaching AI (TK-TAI)	5	M10, M11, M12, M14, M21	.823–.857
4	Disposition Toward Teaching AI (DTAI)	3	M1, M2, M3	.911–.924
Total		26		

Although DTAI retained only three items, this meets the minimum identification requirement for CFA (Kline, 2023). The retained items demonstrated uniformly high SRWs (.911–.924), strong AVE (.840), and high composite reliability (.930), reflecting a coherent construct core: evaluative beliefs about the importance of AI education for students' futures.

The retention of both TPACK-TAI and TK-TAI as separate factors indicates that integrated instructional readiness and narrower technology-specific capability represent empirically distinguishable dimensions within this sample. The HTMT value between the two factors (.864) remained below the .90 discriminant validity threshold (see Table 11), supporting this distinction.

Table 13

Individual Standardized Regression Weights from Confirmatory Factor Analysis

Factor	Item	SRW	Factor	Item	SRW
TPACK-TAI	M17	.763	ATA	M33	.906
TPACK-TAI	M18	.705	ATA	M35	.925
TPACK-TAI	M20	.804	ATA	M36	.902
TPACK-TAI	M22	.834	ATA	M37	.927
TPACK-TAI	M23	.919	ATA	M38	.913
TPACK-TAI	M24	.928	ATA	M40	.897
TPACK-TAI	M25	.945	TK-TAI	M10	.832
TPACK-TAI	M26	.903	TK-TAI	M11	.823
TPACK-TAI	M27	.933	TK-TAI	M12	.857
TPACK-TAI	M28	.893	TK-TAI	M14	.825
TPACK-TAI	M29	.890	TK-TAI	M21	.847
TPACK-TAI	M30	.884	DTAI	M1	.914
			DTAI	M2	.924
			DTAI	M3	.911

Note. All standardized regression weights exceed the .70 retention threshold. Values extracted from IBM SPSS AMOS v26 output.

4.2.2 Configural Invariance

To assess the structural stability of the four-factor model across demographic subgroups, configural invariance was examined by fitting the CFA model separately to gender and school type groups (Table 14).

Table 14

Configural Invariance: Model Fit by Subgroup

Grouping Variable	Group	n	χ²/df	CFI	RMSEA
Gender	Female	134	2.266	.906	.098
Gender	Male	124	2.090	.930	.094
School Type	Public	183	2.384	.931	.087
School Type	Private	75	1.936	.898	.112

Note. RMSEA 90% confidence intervals and SRMR values are not reported for subgroup models because AMOS does not produce stable CI estimates when models are fitted to small samples (n < 200) with complex structures (Kenny et al., 2015). SRMR for the full-sample model was .061 (Table 9).

The four-factor structure demonstrated acceptable fit across both gender groups (CFI ≥ .90, $\chi^2/df < 3.0$), supporting configural invariance for gender. Elevated RMSEA values in the subgroups are attributable to reduced sample sizes (Kenny et al., 2015). For school type, the public school subsample showed acceptable fit; the private school subsample (n = 75) showed marginally below-threshold CFI (.898), likely attributable to the small subsample size. Overall, these results provide preliminary evidence that the four-factor structure holds across the examined subgroups. Full metric and scalar invariance testing with larger subsamples is recommended for future research.

4.3 Reliability Analysis

4.3.1 Item-Total Correlations

Corrected item-total correlations for the 26-item scale ranged from .731 to .920 (Table 15), indicating that all items were strong representatives of their respective subscales.

Table 15

Corrected Item-Total Correlations

Subscale	Item	r	Subscale	Item	r
TPACK-TAI	M17	.780	TK-TAI	M10	.785
TPACK-TAI	M18	.731	TK-TAI	M11	.779
TPACK-TAI	M20	.831	TK-TAI	M12	.824
TPACK-TAI	M22	.850	TK-TAI	M14	.799
TPACK-TAI	M23	.897	TK-TAI	M21	.792
TPACK-TAI	M24	.905	ATA	M33	.888
TPACK-TAI	M25	.920	ATA	M35	.907
TPACK-TAI	M26	.872	ATA	M36	.883
TPACK-TAI	M27	.909	ATA	M37	.911
TPACK-TAI	M28	.871	ATA	M38	.893
TPACK-TAI	M29	.867	ATA	M40	.879
TPACK-TAI	M30	.865	DTAI	M1	.866
			DTAI	M2	.891
			DTAI	M3	.868

4.3.2 Inter-Subscale and Scale-Total Correlations

Table 16

Inter-Subscale and Scale-Total Correlations

TPACK-TAI	TPACK-TAI	ATA	TK-TAI	DTAI	Scale Total
	1.000				**.940**
ATA	.572	1.000			.783
TK-TAI	.818	.509	1.000		.863
DTAI	.482	.762	.440	1.000	.690

Note. All correlations are significant at $p < .01$ (two-tailed).

Inter-subscale correlations (Table 16) ranged from .440 to .818, and subscale-to-total correlations ranged from .690 to .940, confirming that all subscales were significantly and highly correlated with each other and with the overall scale.

4.3.3 Internal Consistency

Table 17

Cronbach's Alpha, Composite Reliability, and McDonald's Omega Values

Scale / Subscale	No. of Items	Cronbach's α	Composite Reliability (CR)	McDonald's ω
Factor 1 — TPACK-TAI	12	.974	.967	.977
Factor 2 — ATA	6	.967	.930	.973
Factor 3 — TK-TAI	5	.921	.954	.941
Factor 4 — DTAI	3	.939	.930	.961
Overall Scale	26	.973	.986	.975

Internal consistency was uniformly strong. All Cronbach's α, CR, and McDonald's ω values (Table 17) exceeded the .70 threshold across both the overall scale and all subscales (Güngör, 2016; Revelle & Zinbarg, 2009; Yaşlıoğlu, 2017). McDonald's omega, which accounts for the congeneric measurement model and does not assume tau-equivalence, confirmed the robustness of these estimates. For TPACK-TAI, the notably high alpha (.974) is accompanied by an inter-item correlation mean of .758 (range = .592–.895), indicating a cohesive factor rather than item redundancy.

5. Discussion

This study developed and validated the Readiness to Teach AI and Behavioral Intention Scale (RTAI-BIS) for K-12 IT teachers. The discussion focuses on what the final structure means theoretically, how the scale relates to prior work, and where its present evidential limits remain.

5.1 Factor Structure and Theoretical Alignment

The final RTAI-BIS comprises 26 items organized into four factors, indicating that readiness to teach AI is best represented as a combination of knowledge-related readiness and intention-related orientation rather than a single disposition. This supports the use of TPACK as a framework for the readiness component, consistent with prior work arguing that AI instruction requires domain-specific integration of content, pedagogy, and technology (Kim et al., 2021a; Yue et al., 2024). The separation between TPACK-TAI and TK-TAI suggests that teachers distinguish between integrated instructional readiness — encompassing explanation, lesson design, assessment, and pedagogical adaptation — and narrower technology-specific capability centered on AI tools, software environments, and coding platforms. This interpretation is consistent with item content: TK-TAI items concern use of specific tools and troubleshooting, whereas TPACK-TAI items emphasize concept explanation, instructional planning, and the coordination of pedagogy with technology. The two factors are theoretically related, as TK is embedded within the broader TPACK framework (Mishra & Koehler, 2006), yet they are empirically non-redundant (HTMT = .864), indicating that tool-level competence and integrated pedagogical-content readiness remain distinguishable in this context. This finding resonates with Velander et al. (2024), who reported that Swedish K-12 teachers' AI content knowledge was largely acquired through incidental exposure and often contained misconceptions, underscoring that technological familiarity and pedagogical-content readiness develop along different trajectories.

The TPB-informed component of the scale is more nuanced. DTAI most closely reflects attitude toward the behavior, because its items concern whether AI education should be provided to students. ATA, by contrast, reflects evaluations of AI as an object or societal phenomenon rather than direct commitment to classroom enactment. The strong ATA-DTAI correlation nevertheless suggests that favorable views of AI and favorable views of teaching AI are empirically connected, even when they are not identical constructs. Habibi et al. (2023) reported a similar pattern: attitudinal beliefs were stronger predictors of preservice teachers' technology integration intentions than knowledge variables alone. In a complementary finding, Addo and Sentance (2023) used self-determination theory to explore K-12 teachers' motivation for teaching AI and found that intrinsic interest in AI, perceived competence, and institutional support were key motivational drivers. These drivers correspond to the ATA and DTAI dimensions of the RTAI-BIS, reinforcing the view that dispositional and evaluative orientations play a central role in shaping teaching behavior.

A central interpretive issue concerns why subjective norms and perceived behavioral control did not survive as independent factors. Subjective norm variance clustered with the dispositional teaching factor, while perceived behavioral control items overlapped with knowledge-related factors — consistent with evidence that normative and control beliefs often merge with attitudinal and capability constructs in teaching contexts (Ajzen, 2002; Armitage & Conner, 2001; Bandura, 1997; Cheung & Cheung Tse, 2021).

This indicates that some TPB components are latently embedded in broader readiness and disposition constructs rather than manifesting as stand-alone subscales.

The shift from five EFA factors to four CFA factors reinforces this interpretation. The removed AI Interest and Engagement factor captured personal engagement with AI tools — behaviors that reflect how often teachers use AI in their own lives rather than whether they are prepared to teach it. Because such personal-use behaviors are likely too transient and idiosyncratic to constitute a stable component of teaching readiness, the factor did not replicate under confirmatory conditions. A response-format difference may also have contributed, as the AIE items used a frequency scale rather than the Likert agreement format of all other subscales (Podsakoff et al., 2003). The retained four-factor model is more tightly aligned with the study's construct focus: teaching-related capability, technology-related capability, attitudinal orientation, and dispositional commitment. The scale therefore captures not general AI enthusiasm but readiness and willingness to teach it.

The transition from five EFA factors to four CFA factors reflects cross-sample refinement rather than fit maximization. The .70 SRW threshold was set a priori, and each removal was evaluated for conceptual as well as statistical justification (see Table 8). Crucially, the AIE dimension was proposed by the EFA sample and independently tested by the CFA sample; its failure to replicate is a genuine cross-validation outcome. That the surviving model produced uniformly high AVE values (.700–.840), clean discriminant validity (all HTMT < .90), and configural invariance across subgroups provides converging evidence against post-hoc overfitting.

5.2 Psychometric Properties

The psychometric results indicate that the final instrument is internally coherent and structurally consistent across independent samples. The EFA and CFA results converged on a structure with interpretable factors, acceptable model fit, and strong item-factor relations. Convergent validity was supported by AVE values above recommended thresholds, and discriminant validity was supported by HTMT values below the conservative .90 criterion (Henseler et al., 2015). The highest HTMT value occurred between TPACK-TAI and TK-TAI, which is theoretically expected because both factors concern knowledge-related readiness; importantly, the value still remained below the threshold, suggesting related but distinguishable constructs.

Reliability evidence was similarly strong. Cronbach's alpha, composite reliability, and McDonald's omega all supported high internal consistency for the total scale and subscales, while item-total correlations indicated that the retained items contributed meaningfully to their respective dimensions. The notably high alpha of TPACK-TAI (.974) warrants comment: with 12 items and corrected item-total correlations ranging from .731 to .920, this value reflects a cohesive factor rather than item redundancy, yet it suggests that a shorter form may be feasible without substantial information loss. Future studies should evaluate whether a reduced TPACK-TAI subscale can maintain comparable validity and reliability. Configural invariance results also suggest that the four-factor model is structurally plausible across gender and school type groups, although stronger subgroup evidence is still needed before broader invariance claims are made.

Given these results, the RTAI-BIS is suitable for research and diagnostic use at the current stage of validation. The retained factor structure appears sufficiently coherent to support substantive interpretation rather than only exploratory description.

Although a dedicated criterion-validation design was not part of the present study, the broader dataset provides limited external-pattern evidence. Teachers who had completed prior AI training scored significantly higher on TPACK-TAI ($t(256) = 4.32$, $p < .001$, $d = 0.54$) and TK-TAI ($t(256) = 5.11$, $p < .001$, $d = 0.64$), with the stronger differentiation on TK-TAI consistent with the expectation that technological knowledge is more directly shaped by formal training. These patterns constitute preliminary known-groups evidence but do not substitute for a formal criterion-validation study based on external performance measures.

5.3 Comparison with Existing Literature

As noted in Section 1, prior work has approached readiness and intention separately. The RTAI-BIS addresses this gap by integrating TPACK-based competence assessment with TPB-informed intention measurement in a single validated instrument (cf. Ayanwale et al., 2022; Chai et al., 2024; Kim et al., 2021a; Yue et al., 2024).

The distinction from adjacent AI-related instruments is equally important. Several recent scales measure teachers' competence in using AI as a pedagogical tool, or their general AI self-efficacy, rather than their readiness to teach AI as curricular content (Celik, 2023; Chiu et al., 2025; Wang & Chuang, 2024). The closest comparable instrument, the TAAI (Guo et al., 2025), assesses teachers' acceptance of AI in education through TAM-based dimensions including perceived usefulness, ease of use, self-efficacy, and anxiety; however, acceptance of AI as an educational resource is conceptually distinct from readiness and intention to teach AI as a subject, which additionally requires pedagogical-content integration and dispositional commitment to classroom enactment. The present scale accordingly addresses this more specific implementation problem: whether teachers are prepared and inclined to teach AI itself. This specificity is important given cross-national evidence that even teachers with moderate AI awareness report feeling unprepared to adapt AI curricula to classroom contexts (Du et al., 2023). In the Turkish context, this is particularly relevant because policy commitments to AI education already exist, but validated tools for assessing teacher preparedness have been lacking (T.C. CBDDO, 2021; T.C. MEB, 2023).

Cross-national studies support the multidimensional nature of the construct. Yau et al. (2023) identified six qualitatively distinct categories in how teachers conceptualize teaching AI, aligning with the RTAI-BIS's separation of knowledge-related and intention-related dimensions. Jatileni et al. (2024) found that Namibian teachers' intention to teach AI depended on attitude and confidence rather than on readiness alone, while Ayanwale and Sanusi (2023) reported significant STEM vs. non-STEM group differences that underscore the importance of domain-specific measurement. Within Turkey, although intention-based scale research has effectively captured teaching readiness for emerging domains (Günbatar & Bakırcı, 2019), no analogous instrument existed for AI — a gap the RTAI-BIS now addresses.

5.4 Practical Implications

Practically, the RTAI-BIS can function as a diagnostic tool for curriculum implementation. School leaders and teacher educators can use it to identify whether weak readiness stems more from pedagogical-content capability, technology-related capability, or attitudinal-dispositional barriers. Consider two contrasting profiles: a teacher scoring high on TPACK-TAI and TK-TAI but low on DTAI has the knowledge to teach AI yet lacks the dispositional commitment to do so — an implementation gap that training alone cannot close. Conversely, a teacher scoring high on DTAI and ATA but low on TPACK-TAI is willing and favorably disposed but lacks the pedagogical-content readiness — a training gap amenable to structured professional development. Teacher education providers can use such profiles to design targeted preservice and in-service support rather than treating AI preparation as a single undifferentiated training need.

5.5 Limitations and Future Directions

No single validation study is fully definitive. Several limitations should be considered. The study was conducted only with IT teachers in Turkey, so the findings cannot yet be generalized across countries, school systems, or teacher populations. Sampling also introduces a concern: non-probability snowball sampling may have introduced selection bias. Relatedly, the instrument relies on self-report data and thus cannot be assumed to reflect actual classroom practice. As reported in Section 4.1.1, converging structural evidence mitigated common method bias concerns; however, future studies should employ procedural remedies such as temporal separation or multi-source data collection (Podsakoff et al., 2003).

Temporal stability was not assessed; future studies should estimate test-retest reliability within a 2–4 week window. The transition from five EFA factors to four CFA factors is both a limitation and a finding: the initial item pool's coverage of personal AI engagement did not survive cross-sample validation, likely because personal tool-use behaviors occupy a different construct space from teaching readiness and because the frequency-scaled items introduced response-format heterogeneity. Future iterations might re-approach the engagement dimension with items anchored to teaching contexts and a consistent response format.

Invariance evidence is preliminary because subgroup sizes, especially for private-school teachers, were limited. On the theoretical side, although the initial pool represented all three TPB components, subjective norms and perceived behavioral control did not remain as independent subscales, limiting the instrument's ability to test the full TPB pathway directly. Finally, the study did not include a dedicated criterion-validation design. The preliminary known-groups evidence reported in Section 5.2 should not be treated as a substitute for criterion-related validation based on classroom observation, lesson quality, or prospective implementation outcomes.

6. Conclusion

This study produced the RTAI-BIS, a four-factor, 26-item instrument that jointly captures TPACK-based readiness and TPB-informed behavioral intention for teaching AI at the K-12 level. The cross-validated factor structure and strong reliability evidence position the scale as a potentially useful diagnostic tool whose multidimensional profile can distinguish knowledge gaps from attitudinal barriers, informing more targeted approaches to teacher preparation rather than uniform training (see Section 5.4 for profiling applications). The availability of Turkish and English versions facilitates future cross-cultural validation and comparison studies as AI education expands internationally.

Knowledge dimensions (TPACK-TAI, TK-TAI) account for the largest share of explained variance, yet attitudinal and dispositional factors remain empirically distinct and necessary — suggesting that teacher readiness for AI education may be knowledge-dominant but not knowledge-sufficient, a pattern that warrants confirmation in future samples.

The findings should be interpreted in light of certain limitations, including the single-country sample, reliance on self-report data, and the absence of test-retest evidence. Several directions deserve priority. Criterion-related validity studies should examine whether RTAI-BIS scores predict observable teaching behaviors and implementation quality. Longitudinal administrations can establish score stability and track how readiness shifts after professional development. Replication across teacher populations outside Turkey and beyond IT teachers will clarify the instrument's generalizability.

As AI education moves from curriculum documents to classroom practice, the gap between intended instruction and teacher capacity will shape what students actually learn. The RTAI-BIS offers one empirically grounded way to measure that gap.

Declarations

Funding: The authors did not receive support from any organization for the submitted work.

Competing Interests: The authors have no relevant financial or non-financial interests to disclose.

Ethics Approval: This study received ethical approval from the Yıldız Technical University Social and Human Sciences Research Ethics Committee (Report No: 20240402851, Decision No: 2024.04; date: April 1, 2024).

Informed Consent: Informed consent was obtained electronically from all participants prior to questionnaire completion. No personally identifiable information was collected.

Data Availability: The anonymized dataset generated during this study is available from the corresponding author upon reasonable request.

Author Contributions: Hasan Tokatlı: Conceptualization, methodology, data collection, formal analysis, and writing – original draft preparation. M. Fatih Erkoç: Supervision, methodological guidance, contribution to data collection, and review & editing. All authors read and approved the final manuscript.

References

1. Addo A (2023) Are you ready to teach AI in schools? Teachers' perspectives of teaching AI in K-12 settings. In Proceedings of the 2023 United Kingdom and Ireland Computing Education Research Conference (pp. 1–7). ACM. https://doi.org/10.1145/3610969.3610973
2. Addo A, Sentance S (2023) Teachers' motivation for teaching AI in K-12 settings. In Proceedings of the 2023 Conference on Human Centered Artificial Intelligence: Education and Practice (pp. 1–5). ACM. https://doi.org/10.1145/3633083.3633192
3. Ajzen I (1991) The theory of planned behavior. Organ Behav Hum Decis Process 50(2):179–211. https://doi.org/10.1016/0749-5978(91)90020-T
4. Ajzen I (2002) Perceived behavioral control, self-efficacy, locus of control, and the theory of planned behavior. J Appl Soc Psychol 32(4):665–683. https://doi.org/10.1111/j.1559-1816.2002.tb00236.x
5. Ajzen I (2011) The theory of planned behaviour: Reactions and reflections. Psychol Health 26(9):1113–1127. https://doi.org/10.1080/08870446.2011.613995
6. Ajzen I (2020) The theory of planned behavior: Frequently asked questions. Hum Behav Emerg Technol 2(4):314–324. https://doi.org/10.1002/hbe2.195
7. Armitage CJ, Conner M (2001) Efficacy of the theory of planned behaviour: A meta-analytic review. Br J Soc Psychol 40(4):471–499. https://doi.org/10.1348/014466601164939
8. Ayanwale MA, Sanusi IT (2023) Perceptions of STEM vs. non-STEM teachers toward teaching artificial intelligence. In 2023 IEEE AFRICON (pp. 1–5). IEEE. https://doi.org/10.1109/africon55910.2023.10293455
9. Ayanwale MA, Sanusi IT, Adelana OP, Aruleba KD, Oyelere SS (2022) Teachers' readiness and intention to teach artificial intelligence in schools. Computers Education: Artif Intell 3:100099. https://doi.org/10.1016/j.caeai.2022.100099
10. Bandura A (1997) Self-efficacy: The exercise of control. W. H. Freeman
11. Beauducel A, Herzberg PY (2006) On the performance of maximum likelihood versus means and variance adjusted weighted least squares estimation in CFA. Struct Equ Model 13(2):186–203. https://doi.org/10.1207/s15328007sem1302_2
12. Bessen JE (2018) AI and jobs: The role of demand (NBER Working Paper No. 24235). National Bureau of Economic Research. https://doi.org/10.3386/w24235
13. Bollen KA, Stine RA (1992) Bootstrapping goodness-of-fit measures in structural equation models. Sociol Methods Res 21(2):205–229. https://doi.org/10.1177/0049124192021002004
14. Büyüköztürk Ş (2002) Faktör analizi: Temel kavramlar ve ölçek geliştirmede kullanımı. Kuram ve Uygulamada Eğitim Yönetimi 32:470–483

15. Casal-Otero L, Catala A, Fernández-Morante C, Taboada M, Cebreiro B, Barro S (2023) AI literacy in K-12: A systematic literature review. Int J STEM Educ 10(1):29. https://doi.org/10.1186/s40594-023-00418-7
16. Celik I (2023) Towards Intelligent-TPACK: An empirical study on teachers' professional knowledge to ethically integrate artificial intelligence (AI)-based tools into education. Comput Hum Behav 138:107468. https://doi.org/10.1016/j.chb.2022.107468
17. Chai CS, Liang S, Wang X (2024) A survey study of Chinese teachers' continuous intentions to teach artificial intelligence. Educ Inform Technol 29(11):14015–14034. https://doi.org/10.1007/s10639-023-12430-z
18. Cheung HC, Cheung Tse AW (2021) Hong Kong science in-service teachers' behavioural intention towards STEM education and their technological pedagogical content knowledge (TPACK). In 2021 IEEE International Conference on Engineering, Technology & Education (TALE) (pp. 630–637). IEEE. https://doi.org/10.1109/TALE52509.2021.9678933
19. Chiu TKF, Ahmad Z, Çoban M (2025) Development and validation of teacher artificial intelligence (AI) competence self-efficacy (TAICS) scale. Educ Inform Technol 30(5):6667–6685. https://doi.org/10.1007/s10639-024-13094-z
20. Costello AB, Osborne JW (2005) Best practices in exploratory factor analysis: Four recommendations for getting the most from your analysis. Practical Assess Res Evaluation 10(7):1–9. https://doi.org/10.7275/jyj1-4868
21. Davis FD (1989) Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Q 13(3):319–340. https://doi.org/10.2307/249008
22. DeVellis RF (2017) Scale development: Theory and applications, 4th edn. Sage
23. Di Battista A, Grayling S, Hasselaar E, Leopold T, Li R, Rayner M, Zahidi S (2023) Future of Jobs Report 2023. World Economic Forum
24. Du X, Taylor M, Blumofe N (2023) Exploring teachers' self-perceived effectiveness, readiness, and ability to adapt an AI literacy curriculum in an international classroom context. In Proceedings of the 2023 AERA Annual Meeting. AERA. https://doi.org/10.3102/2004458
25. Erkuş A (2019) Psikolojide ölçme ve ölçek geliştirme. Pegem Akademi. 4th edn.https://doi.org/10.14527/9786053643111
26. Fabrigar LR, Wegener DT, MacCallum RC, Strahan EJ (1999) Evaluating the use of exploratory factor analysis in psychological research. Psychol Methods 4(3):272–299. https://doi.org/10.1037/1082-989X.4.3.272
27. Fornell C, Larcker DF (1981) Evaluating structural equation models with unobservable variables and measurement error. J Mark Res 18(1):39–50. https://doi.org/10.1177/002224378101800104
28. Frey CB, Osborne MA (2017) The future of employment: How susceptible are jobs to computerisation? Technol Forecast Soc Chang 114:254–280. https://doi.org/10.1016/j.techfore.2016.08.019

29. Günbatar MS, Bakırcı H (2019) STEM teaching intention and computational thinking skills of pre-service teachers. Educ Inform Technol 24(2):1615–1629. https://doi.org/10.1007/s10639-018-9849-5
30. Güngör D (2016) Psikolojide ölçme araçlarının geliştirilmesi ve uyarlanması kılavuzu. Türk Psikoloji Yazıları 19(38):104–112
31. Guo S, Shi L, Zhai X (2025) Developing and validating an instrument for teachers' acceptance of artificial intelligence in education. Educ Inform Technol 30:13439–13461. https://doi.org/10.1007/s10639-025-13338-6
32. Habibi A, Riady Y, Al-Adwan S, A., Albelbisi NA (2023) Beliefs and knowledge for pre-service teachers' technology integration during teaching practice: An extended theory of planned behavior. Computers Schools 40(2):107–132. https://doi.org/10.1080/07380569.2022.2124752
33. Hair JF, Black WC, Babin BJ, Anderson RE (2019) Multivariate data analysis (8th ed.). Cengage
34. Henseler J, Ringle CM, Sarstedt M (2015) A new criterion for assessing discriminant validity in variance-based structural equation modeling. J Acad Mark Sci 43(1):115–135. https://doi.org/10.1007/s11747-014-0403-8
35. Hu L, Bentler PM (1999) Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. Struct Equ Model 6(1):1–55. https://doi.org/10.1080/10705519909540118
36. International Test Commission (2018) The ITC guidelines for translating and adapting tests. Int J Test 18(2):101–134. 2nd edn.https://doi.org/10.1080/15305058.2017.1398166
37. Jatileni CN, Sanusi IT, Olaleye SA, Ayanwale MA, Agbo FJ, Oyelere PB (2024) Artificial intelligence in compulsory level of education: Perspectives from Namibian in-service teachers. Educ Inform Technol 29(10):12569–12596. https://doi.org/10.1007/s10639-023-12341-z
38. Kenny DA, Kaniskan B, McCoach DB (2015) The performance of RMSEA in models with small degrees of freedom. Sociol Methods Res 44(3):486–507. https://doi.org/10.1177/0049124114543236
39. Kim S, Jang Y, Choi S, Kim W, Jung H, Kim S, Kim H (2021a) Analyzing teacher competency with TPACK for K-12 AI education. KI - Künstliche Intelligenz 35(2):139–151. https://doi.org/10.1007/s13218-021-00731-9
40. Kim S, Jang Y, Kim W, Choi S, Jung H, Kim S, Kim H (2021b) Why and what to teach: AI curriculum for elementary school. Proceedings of the AAAI Conference on Artificial Intelligence, 35, 15569–15576. https://doi.org/10.1609/aaai.v35i17.17833
41. Kline RB (2023) Principles and practice of structural equation modeling, 5th edn. Guilford Press
42. Koehler MJ, Mishra P (2009) What is technological pedagogical content knowledge? Contemp Issues Technol Teacher Educ 9(1):60–70
43. Koyuncu İ, Kılıç AF (2019) Açımlayıcı ve doğrulayıcı faktör analizlerinin kullanımı: Bir doküman incelemesi. Eğitim ve Bilim 44(198):361–388. https://doi.org/10.15390/eb.2019.7665

44. Li CH (2016) Confirmatory factor analysis with ordinal data: Comparing robust maximum likelihood and diagonally weighted least squares. Behav Res Methods 48:936–949. https://doi.org/10.3758/s13428-015-0619-7
45. Lin K-Y, Williams PJ (2016) Taiwanese preservice teachers' science, technology, engineering, and mathematics teaching intention. Int J Sci Math Educ 14(6):1021–1036. https://doi.org/10.1007/s10763-015-9645-2
46. 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). ACM. https://doi.org/10.1145/3313831.3376727
47. Lorenzo-Seva U, ten Berge JMF (2006) Tucker's congruence coefficient as a meaningful index of factor similarity. Methodology 2(2):57–64. https://doi.org/10.1027/1614-2241.2.2.57
48. Martinez W (2018) How science and technology developments impact employment and education. Proceedings of the National Academy of Sciences, 115(50), 12624–12629. https://doi.org/10.1073/pnas.1803216115
49. Mishra P, Koehler MJ (2006) Technological pedagogical content knowledge: A framework for teacher knowledge. Teachers Coll Record 108(6):1017–1054. https://doi.org/10.1177/016146810610800610
50. Ng DTK, Leung JKL, Su J, Ng RCW, Chu SKW (2023) Teachers' AI digital competencies and twenty-first century skills in the post-pandemic world. Education Tech Research Dev 71:137–161. https://doi.org/10.1007/s11423-023-10203-6
51. Nunnally JC, Bernstein IH (1994) Psychometric theory, 3rd edn. McGraw-Hill
52. Panth B, Maclean R (eds) (2020) Anticipating and preparing for emerging skills and jobs. Springer. https://doi.org/10.1007/978-981-15-7018-6
53. Podsakoff PM, MacKenzie SB, Lee J-Y, Podsakoff NP (2003) Common method biases in behavioral research: A critical review of the literature and recommended remedies. J Appl Psychol 88(5):879–903. https://doi.org/10.1037/0021-9010.88.5.879
54. Ramazanoğlu M, Akın T (2025) AI readiness scale for teachers: Development and validation. Educ Inform Technol 30:6869–6897. https://doi.org/10.1007/s10639-024-13087-y
55. Revelle W, Zinbarg RE (2009) Coefficients alpha, beta, omega, and the glb: Comments on Sijtsma. Psychometrika 74(1):145–154. https://doi.org/10.1007/s11336-008-9102-z
56. Rhemtulla M, Brosseau-Liard PÉ, Savalei V (2012) When can categorical variables be treated as continuous? A comparison of robust continuous and categorical SEM estimation methods under suboptimal conditions. Psychol Methods 17(3):354–373. https://doi.org/10.1037/a0029315
57. Rizvi S, Waite J, Sentance S (2023) Artificial intelligence teaching and learning in K-12 from 2019 to 2022: A systematic literature review. Computers Education: Artif Intell 4:100145. https://doi.org/10.1016/j.caeai.2023.100145
58. Sanusi IT, Ayanwale MA, Chiu TKF (2024) Investigating the moderating effects of social good and confidence on teachers' intention to prepare school students for artificial intelligence education.

Educ Inform Technol 29(1):273–295. https://doi.org/10.1007/s10639-023-12250-1

59. Sanusi IT, Oyelere SS, Vartiainen H, Suhonen J, Tukiainen M (2023) A systematic review of teaching and learning machine learning in K-12 education. Educ Inform Technol 28:5967–5997. https://doi.org/10.1007/s10639-022-11416-7
60. Scherer R, Siddiq F, Tondeur J (2019) The technology acceptance model (TAM): A meta-analytic structural equation modeling approach to explaining teachers' adoption of digital technology in education. Comput Educ 128:13–35. https://doi.org/10.1016/j.compedu.2018.09.009
61. Sharma S, Mukherjee S, Kumar A, Dillon WR (2005) A simulation study to investigate the use of cutoff values for assessing model fit in covariance structure models. J Bus Res 58(7):935–943. https://doi.org/10.1016/j.jbusres.2003.10.007
62. Shulman LS (1986) Those who understand: Knowledge growth in teaching. Educational Researcher 15(2):4–14. https://doi.org/10.2307/1175860
63. Shulman LS (1987) Knowledge and teaching: Foundations of the new reform. Harv Educational Rev 57(1):1–22. https://doi.org/10.17763/haer.57.1.j463w79r56455411
64. Su J, Guo K, Chen X, Chu SKW (2024) Teaching artificial intelligence in K-12 classrooms: A scoping review. Interact Learn Environ 32(9):5207–5226. https://doi.org/10.1080/10494820.2023.2212706
65. Su J, Zhong Y, Ng DTK (2022) A meta-review of literature on educational approaches for teaching AI at the K-12 levels in the Asia-Pacific region. Computers Education: Artif Intell 3:100065. https://doi.org/10.1016/j.caeai.2022.100065
66. T.C. Cumhurbaşkanlığı Dijital Dönüşüm Ofisi [CBDDO]. (2021) Ulusal yapay zeka stratejisi 2021–2025. https://cbddo.gov.tr/SharedFolderServer/Genel/File/TR-UlusalYZStratejisi2021-2025.pdf
67. T.C. Millî Eğitim Bakanlığı [MEB]. (2023) Yapay zekâ uygulamaları dersi öğretim programı. https://mufredat.meb.gov.tr/Dosyalar/2023112493011132-23174117_yapayzekauygulamalaridersiogretimprogrami_3.23.pdf
68. Touretzky DS, Gardner-McCune C, Martin F, Seehorn D (2019) Envisioning AI for K-12: What should every child know about AI? Proceedings of the AAAI Conference on Artificial Intelligence, 33(1), 9795–9799. https://doi.org/10.1609/aaai.v33i01.33019795
69. UNESCO (2021) AI and education: Guidance for policy-makers. UNESCO Publishing. https://doi.org/10.54675/PCSP7350
70. Velander J, Taiye MA, Otero N, Milrad M (2024) Artificial intelligence in K-12 education: Eliciting and reflecting on Swedish teachers' understanding of AI and its implications for teaching & learning. Educ Inform Technol 29(4):4085–4105. https://doi.org/10.1007/s10639-023-11990-4
71. Venkatesh V, Morris MG, Davis GB, Davis FD (2003) User acceptance of information technology: Toward a unified view. MIS Q 27(3):425–478. https://doi.org/10.2307/30036540
72. Wang Y-Y, Chuang Y-W (2024) Artificial intelligence self-efficacy: Scale development and validation. Educ Inform Technol 29:4535–4561. https://doi.org/10.1007/s10639-023-12015-w

73. Yaşlıoğlu MM (2017) Sosyal bilimlerde faktör analizi ve geçerlilik: Keşfedici ve doğrulayıcı faktör analizlerinin kullanılması. İstanbul Üniversitesi İşletme Fakültesi Dergisi, 46(özel sayı), 74–85
74. Yau KW, Chai CS, Chiu TKF, Meng H, King I, Yam Y (2023) A phenomenographic approach on teacher conceptions of teaching artificial intelligence (AI) in K-12 schools. Educ Inform Technol 28(1):1041–1064. https://doi.org/10.1007/s10639-022-11161-x
75. Yue M, Jong MSY, Ng DTK (2024) Understanding K-12 teachers' technological pedagogical content knowledge readiness and attitudes toward artificial intelligence education. Educ Inform Technol 29(13):16951–16976. https://doi.org/10.1007/s10639-024-12621-2

Figures

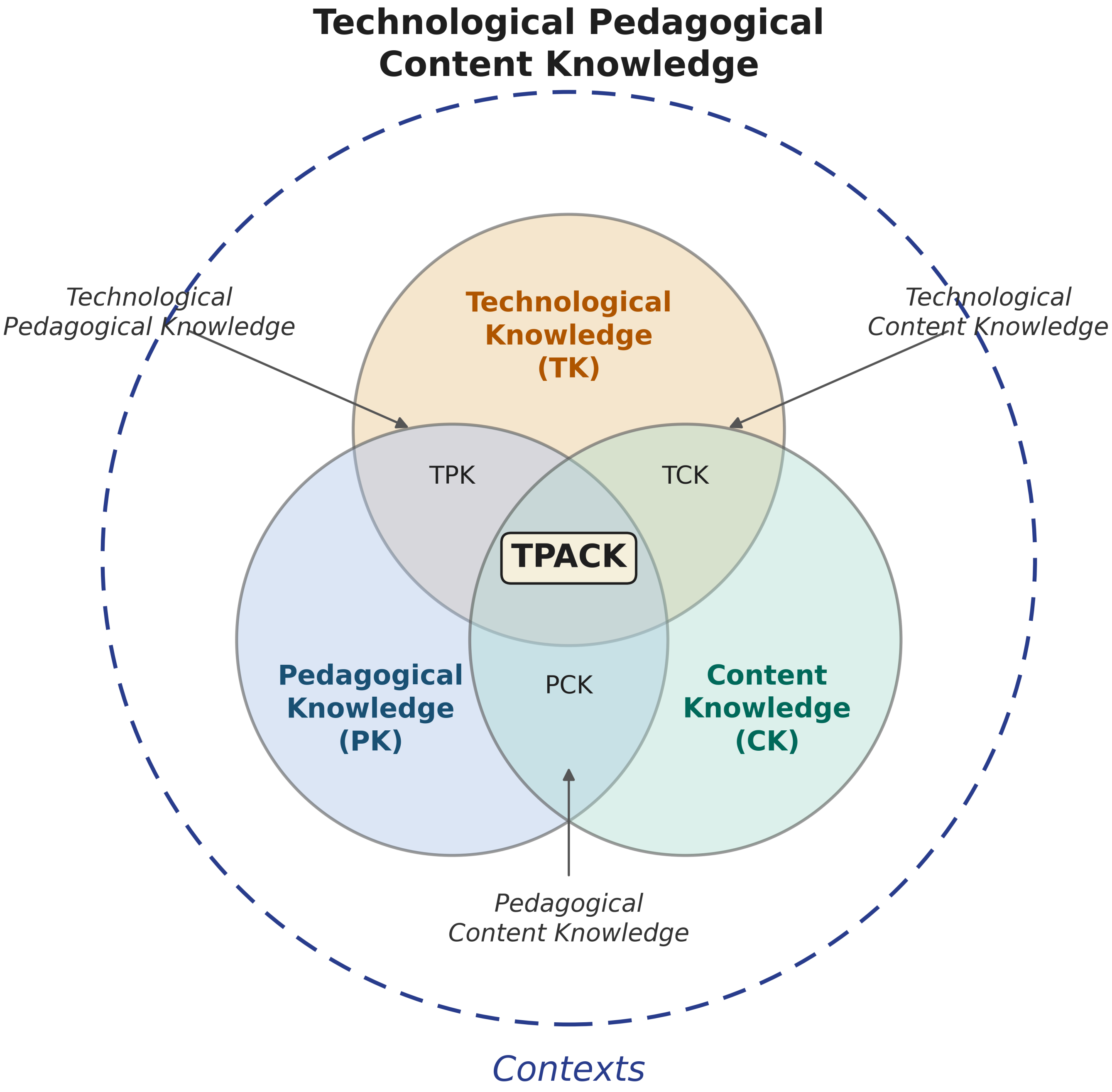

Figure 1

The TPACK framework and its knowledge components (adapted from Koehler & Mishra, 2009). TK = Technological Knowledge; PK = Pedagogical Knowledge; CK = Content Knowledge; TPK = Technological Pedagogical Knowledge; TCK = Technological Content Knowledge; PCK = Pedagogical Content Knowledge; TPACK = Technological Pedagogical Content Knowledge. Reproduced by permission of the publisher, © 2012 by tpack.org

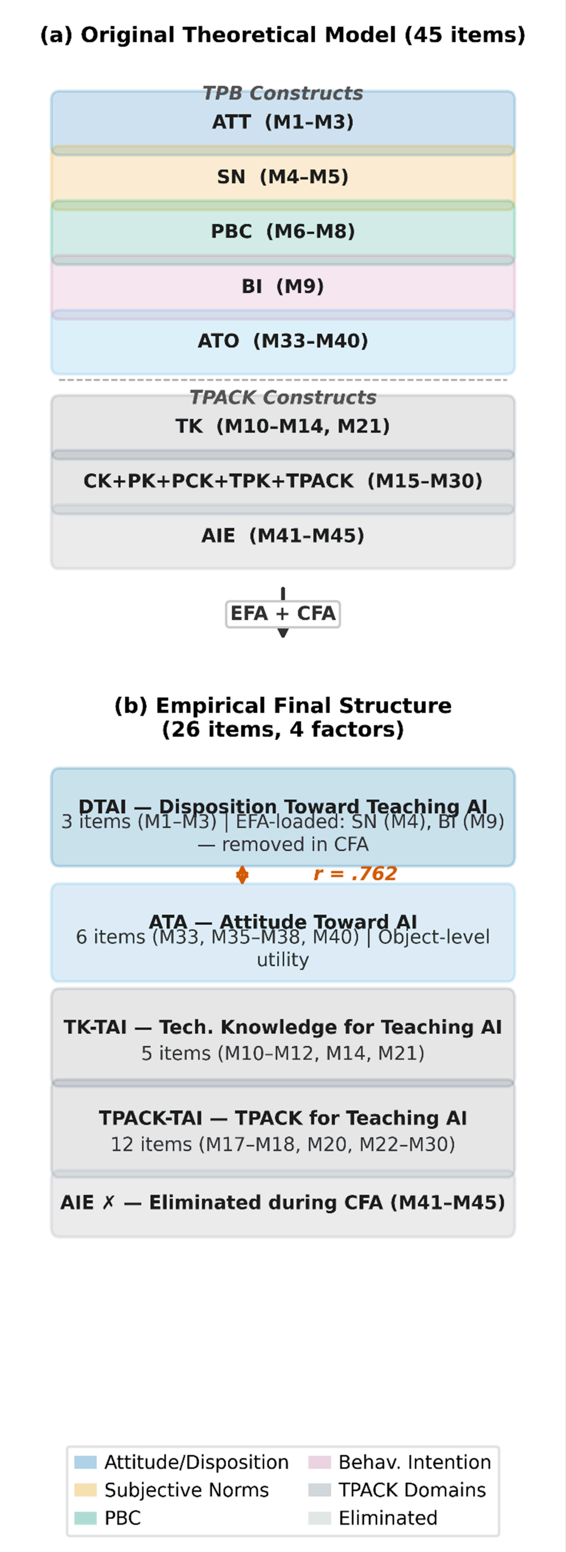

Figure 2

Theoretical-to-empirical mapping of TPB and TPACK constructs in the RTAI-BIS. Panel (a) shows the original 45-item candidate pool with all three TPB components represented. Panel (b) shows the final 26-item structure after EFA/CFA refinement. Subjective norm (M4) and behavioral intention (M9) items loaded on the DTAI factor during EFA but subsequently removed in CFA for failing the SRW threshold. PBC items (M6–M8) were eliminated during EFA for failing loading criteria

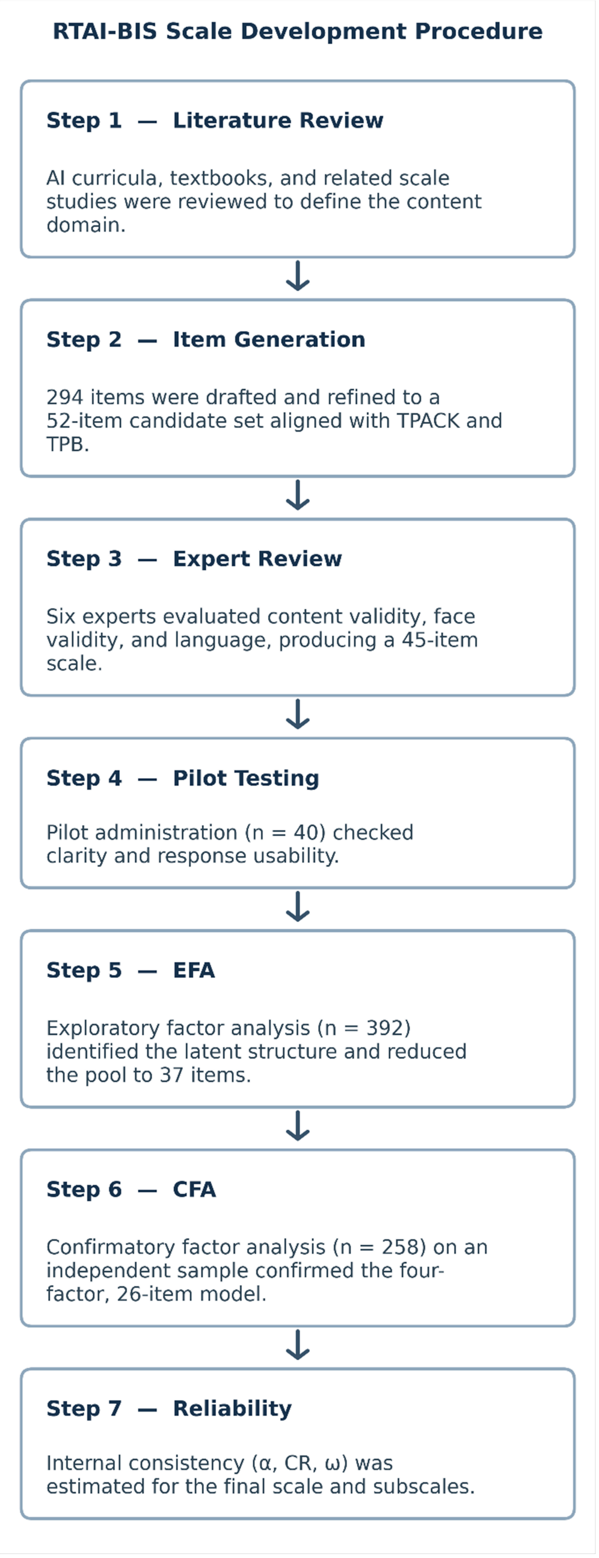

Figure 3

Scale development procedure for the RTAI-BIS, from literature review and item generation through expert review, pilot testing, exploratory and confirmatory factor analyses, and reliability assessment

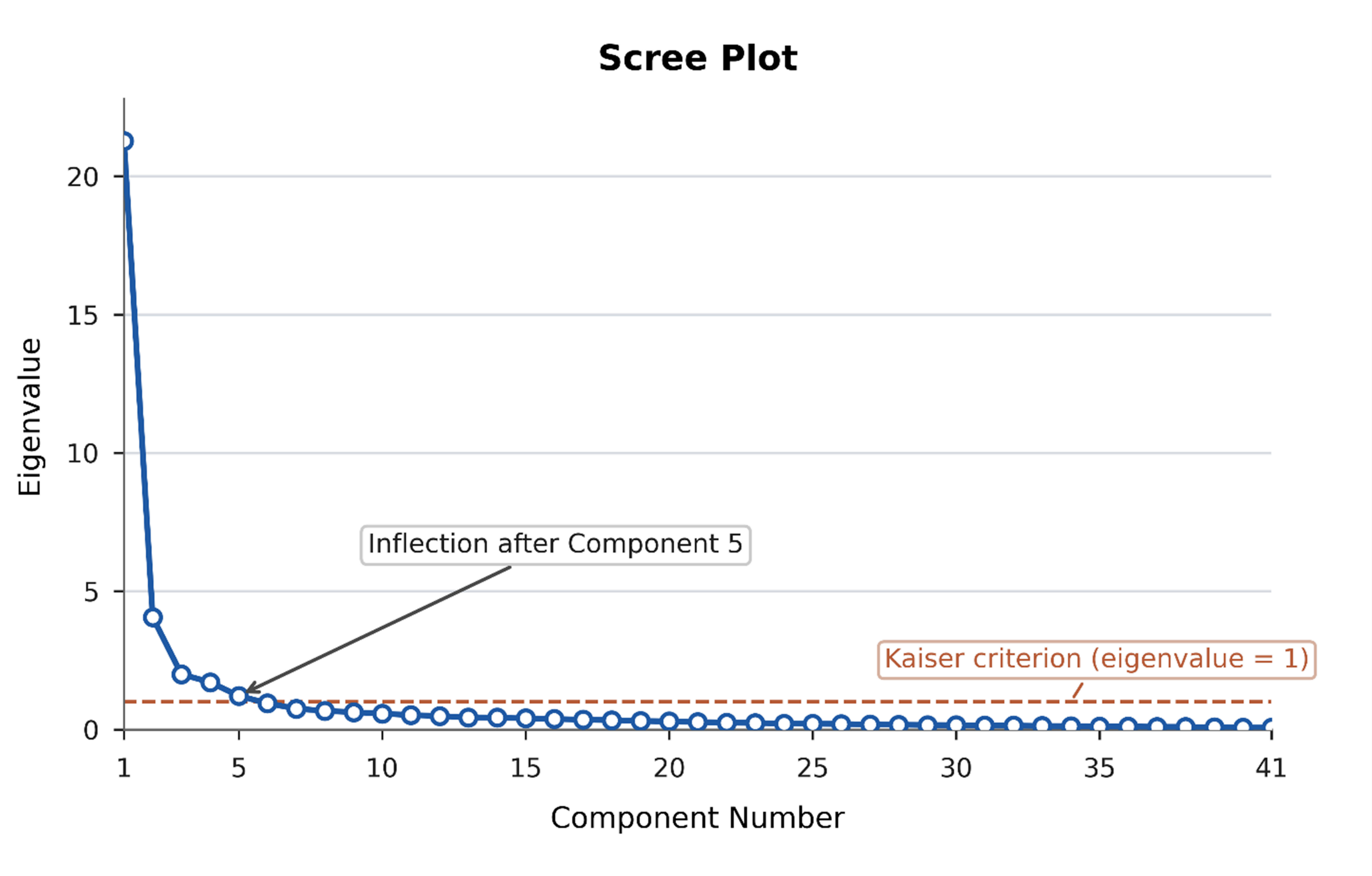

Figure 4

Scree plot for the 41-item EFA. The arrow indicates the inflection point after the fifth component, supporting a five-factor solution

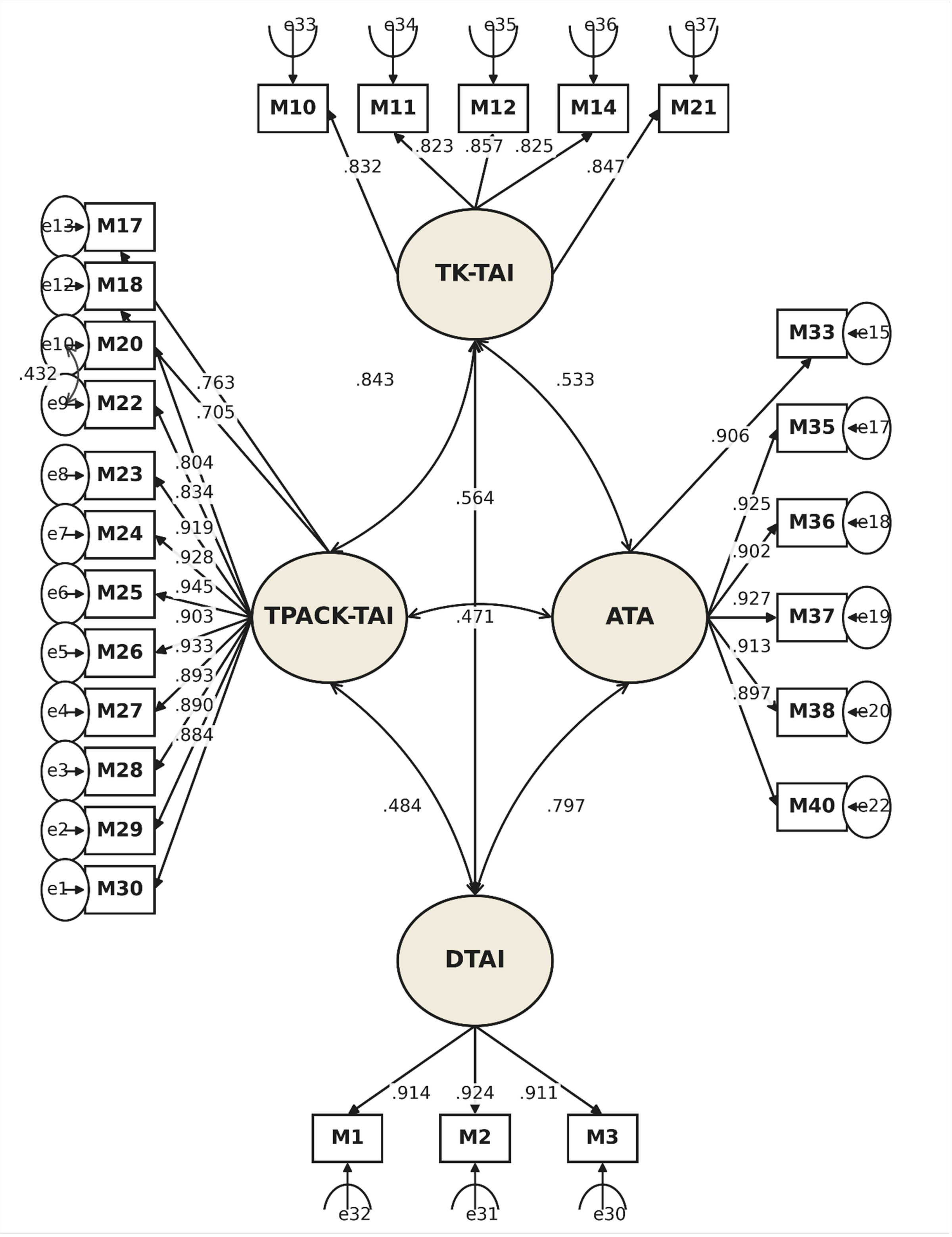

Figure 5

Confirmatory factor analysis path diagram for the four-factor, 26-item RTAI-BIS model. Standardized regression weights are displayed on paths. Inter-factor correlations shown are CFA standardized latent estimates. Error covariance between e9 and e10 (M20 ↔ M22) is specified

Supplementary Files