

Validation of the adapted attitudes toward online teaching scale for Iranian EFL teachers: A Rasch model analysis

Neda Kianinezhad ¹

¹ Islamic Azad University, Mashhad, Iran

ARTICLE HISTORY

Received: 23.01.2024

Accepted: 19.03.2024

KEYWORDS

Attitudes toward online teaching

Scale adaptation

Iranian EFL teachers

Rasch model analysis

Validity

CORRESPONDENCE

Neda Kianinezhad,
kianinezhad.neda@gmail.com

ABSTRACT

This study addresses a critical gap by adapting and validating an instrument aimed at evaluating the perspectives of Iranian English as a Foreign Language (EFL) teachers in online teaching. Leveraging 'Attitude Toward Online Teaching' scale, the survey involved 160 Iranian EFL teachers. The precise adaptation and analysis methods, centered on the Rasch measurement model, substantially bolstered the scale's analytical rigor. Subsequently, through translation and validation procedures, coupled with Rasch models, affirmed the scale's robust reliability and validity. Establishing validity evidence for the translated scale involved data fitting to the Rasch model, substantiating the cross-cultural validity of both the construct and the Persian test version. Participant responses were pivotal in identifying and eliminating items that did not align well, resulting in a more refined 24-item version which indicates unidimensionality of the scale. Additionally, the discernible differentiation among the scale's categories, evident in the sequence of category thresholds in participant responses, underscores their substantial importance. The consistently high reliability observed across measurement validates the scale's ability to accurately capture attitudes within the Iranian EFL context. This polished Persian adaptation of the instrument emerges as a transformative asset, holding implications for pedagogical practices, curriculum development, and teacher training in online EFL education. It marks a significant stride in advancing EFL education in the digital era, introducing a culturally contextualized measurement tool tailored explicitly to the Iranian EFL setting. Lastly, insights empower stakeholders to refine education strategies, advancing the digital educational landscape.

Introduction

Education, a fundamental pillar of personal development, continually evolves in response to technological advancements throughout an individual's lifespan (Sungur & Ateş, 2023; Jovanova-Mitkovska, 2010; Ghavifekr & Rosdy, 2015; Raja & Nagasubramani, 2018; Dos Santos, 2024). The pervasive influence of the internet globally has radically reshaped education, bolstering pedagogical efficacy through the seamless integration of technology (Lateef & Alaba, 2013). That

is, the internet has significantly transformed education, enhancing teaching efficiency through technological integration. In today's educational panorama, educators, particularly tech-savvy millennials, find themselves at a crucial juncture to enrich their digital competencies (Kalanda, 2005). The past decade has witnessed substantial strides in the online domain of higher education, fervently advocated by Philip et al. (2009). Yet, a divergence persists among educators concerning the efficacy of online instruction: while some view it as a remedy for accessibility barriers, others emphasize the irreplaceability of traditional pedagogical methods (Enayati et al., 2012). The prevailing approach inclines towards amalgamating technology with traditional methods to augment student outcomes (Salmon, 2012; Costley, 2014), underscoring the escalating significance of technology and digital literacy in education.

Blended learning and online education seamlessly became integrated into higher education in the early 21st century (Singh & Thurman, 2019). The global disruptions caused by the 2019 pandemic further propelled the adoption of online or synchronous learning in educational institutions (Sahoo, 2020). This sudden shift, i.e., the necessity to swiftly transition to online learning, necessitated a rapid evaluation of educators' preparedness and perspectives towards online teaching, given the constrained timeframe for comprehensive planning (Sahoo, 2020; Ng et al., 2023).

Educators' attitudes wield substantial influence over their readiness for the digital realm in higher education (Buhl-Wiggers et al., 2023; Martin et al., 2019; Trivedi et al., 2018). The effective integration of information and communication technology (ICT) into classrooms hinges upon educators' cognitive frameworks, beliefs, and attitudes towards ICT (Sang et al., 2010). Likewise, the success of an online course intertwines with elements such as the learning environment, task assignments, and students' attitudes towards ICT (Wasserman & Migdal, 2019).

Facilitating equitable access to quality education demands an exploration of factors shaping educators' acceptance of online teaching (Hung, 2016; Kebritchi et al., 2017). Consequently, educators' attitudes towards online education profoundly impact their performance in digital pedagogy (Wasserman & Migdal, 2019). Importantly, the attitudes of English as a Foreign Language (EFL) teachers towards online teaching have gained prominence in the evolving language education landscape (Ng et al., 2023). Given the rapid technological advancements and the global pivot towards online education, comprehending EFL teachers' perspectives, beliefs, and experiences in online language instruction becomes imperative for adept implementation and continual refinement of digital teaching methodologies. Recent research underscores the role of online platforms in fostering learner autonomy and self-directed learning among EFL students (Stickler et al., 2021; Yosintha & Yuniarti, 2021). These insights offer valuable guidance on how EFL educators can leverage online learning environments to nurture learner autonomy and empower students to steer their language learning journey.

In fact, attitude, a multidimensional psychological construct, serves as a foundational framework for comprehending diverse phenomena. It encapsulates the ability to discern and evaluate experiences, intellect, emotions, and consequent behaviors (Schwarz, 2007). Besides, attitudes encapsulate individual perspectives and personal attributions of significance (Krosnick & Petty, 1995). Amidst the adoption and transformation propelled by technology, attitude assumes a pivotal role (Krishnakumar & Rajesh, 2011).

Educators' attitudes towards online education are shaped by factors such as knowledge, willingness to learn (Papp, 1998), comfort levels (Nair & Das, 2012), beliefs, and the external environment. Additionally, cultural norms, values, and societal conditions shape individual mindsets (Gardner et al., 1993). Educators exhibiting positive attitudes often demonstrate proficiency in adapting to new technologies (Uzunboylyu, 2007). Despite reports of favorable

attitudes towards e-learning among educators (Suri & Sharma, 2017; Akaslan & Law, 2011; Khukalenko et al., 2022), the practical assimilation of technology in classrooms remains sporadic.

To ensure the widespread applicability of the 'Attitudes Toward Online Teaching' scale across diverse cultural landscapes, validating its reliability and validity across various populations becomes imperative. While the scale has demonstrated commendable psychometric properties and construct validity within an English context, assessing its cross-cultural validity is crucial for a more thorough understanding. Notably, no identified investigations within the Iranian context have specifically addressed the validation of the 'Attitudes Toward Online Teaching' (ATOT) scale using the Rasch model (Rasch, 1960/ 1980). This research gap is significant due to the unique cultural and educational nuances in Iran that may influence attitudes toward online teaching differently than in English-speaking contexts. To address this gap, the study focuses on validating the Persian adaptation of the ATOT scale among Iranian EFL teachers, utilizing the Rasch measurement model (Rasch, 1960/ 1980).

Further, crafted by the Danish mathematician and statistician Georg Rasch, this model serves as a revered analytical tool in the social sciences, meticulously evaluating measurement instrument data and probing test and construct validity. It is acknowledged for attributes like invariance and interval scaling, contingent on the foundational assumption of unidimensionality, i.e., the assumption that the underlying trait being measured is a single, unidimensional construct (Kazemi et al., 2020; Baghaei, 2009). The Rasch measurement model harbors the potential to reveal distinctive attributes, facilitating seamless alignment between data and model constructs. For instance, it ensures invariance across different groups or conditions and provides interval-scaled measurements, enhancing its utility in various research applications within the social sciences. Accordingly, this study aims to rigorously examine and validate the 'Attitudes Toward Online Teaching' scale in the context of Iran, utilizing the Rasch model, contributing to a more comprehensive understanding of online teaching attitudes across diverse cultural and linguistic backgrounds.

Method

Participants and setting

In this study, a meticulously selected cohort of 160 Iranian EFL teachers actively engaged, comprising 58 males (36.25%) and 102 females (63.75%). The determination of the sample size closely adhered to Linacre's guidelines (1994), ensuring its suitability for Rasch model analysis. This diverse participant ensemble exhibited pronounced diversity across multiple dimensions, encompassing gender representation, educational attainment spanning from AA to Ph.D. levels, and a spectrum of workplace settings, including schools, institutes, universities, and various other contexts; moreover, participants showcased varied teaching experiences, tenures in English Language Teaching, and a wide age demographic. Participant selection employed a convenient sampling method utilizing Google Forms. It is crucial to emphasize that all participants were native Persian speakers and held Iranian citizenship. Stringent adherence to ethical guidelines was upheld to safeguard the confidentiality and anonymity of all individuals involved. This rigorous approach aimed to bolster the representativeness of the participant cohort, thereby enhancing the reliability and credibility of the amassed data.

Measure

The integration of the Persian adaptation of the "Attitudes Toward Online Teaching" scale in this study marks a significant stride, providing profound insights into the perspectives of Iranian

English as a Foreign Language (EFL) teachers regarding online education. Originally formulated and validated in English by Sangwan et al. (2021), the scale underwent meticulous translation into Persian, yielding a comprehensive set of 29 crafted Likert scale items. The survey comprised two distinct segments. The first section aimed to gather demographic information from participants, including gender, educational background, workplace, teaching experience, and age. The subsequent section involved participants in evaluating content areas relevant to the study, utilizing a refined five-point Likert scale. Participants were tasked with carefully assessing each item on a scale from 1 (strongly disagree) to 5 (strongly agree), facilitating a comprehensive exploration of instructors' perspectives on online teaching. To bolster the scholarly rigor of the study, the scale's reliability and validity underwent rigorous scrutiny through Rasch model analysis. This rigorous examination has enriched our understanding of the psychometric properties inherent in the Persian "Attitudes Towards Online Teaching" scale, particularly within the unique context of Iranian EFL educators. Such a scrupulous approach to psychometric analysis underscores the robustness of the research findings, further solidifying their scholarly significance. No item required reverse scoring, and higher scores indicated higher levels of positive attitudes toward online teaching.

Adaptation procedure

In its pursuit of examining the cultural relevance and accuracy in measuring attitudes toward online teaching among Iranian EFL teachers, this study rigorously adhered to a systematic procedure. To achieve this objective, the research utilized a questionnaire initially developed by Sangwan et al. (2021) and subjected it to a thorough process of translation and validation. For this purpose, the study actively involved experts in educational technology and educational psychology. Firstly, the questionnaire was translated from English to Persian in a forward translation phase (Khodabakhshzadeh & Shoahosseini, 2021). Subsequently, a rigorous back-translation process was undertaken to ensure linguistic precision and cultural authenticity and discrepancies were resolved (Moradi, 2020; Tabatabaee-Yazdi & Baghaei 2022). Note that the experts evaluated the translations individually to provide their insights and suggestions. This step-by-step approach ensured a comprehensive adaptation of the instrument, aligning it effectively with the specific context of Iranian EFL teachers.

Findings

Preliminary analyses

In the initial stage of analysis, preliminary analyses were conducted. Descriptive statistics for the scale data unveiled a standard deviation of 21.39 and a mean of 88.73. Notably, both skewness and kurtosis values were found to fall within the accepted range (± 2), suggesting a normal distribution of the data. Moreover, an analysis of item-total statistics for the scale was performed. Remarkably, all values of the corrected item-total correlation surpassed 0.30, indicating a robust consistency in measuring the same variable across items.

Validation

In the realm of education and social sciences, various statistical methodologies play a pivotal role in formulating, evaluating, validating, and interpreting test outcomes (e.g., Asrijanty, 2014; Abdullah et al., 2021; Baghaei et al., 2017; Borsboom et al., 2004; Kianinezhad, 2024; Khajavy et al., 2018; Lambri et al., 2019; Nadri et al., 2019; Ningsih et al., 2021; Sarabi, 2017; Saghafi et al., 2021; Shirvana et al., 2016).

Among these methodologies, Rasch models stand out significantly as a well-established approach for validating questionnaires and tests within these disciplines, i.e., classified as latent variable

models. Rasch models elucidate concealed traits by revealing an underlying factor that connects observed variables when data adhere to the Rasch model (Andrich, 1988; Baghaei & Shoahosseini, 2019). This latent trait profoundly influences how individuals respond to test items, providing crucial insights into consistent answering patterns and holding fundamental importance within the validity framework (Borsboom et al., 2004).

Moreover, Rasch models present distinct advantages over Confirmatory Factor Analysis (CFA). These models maintain robust invariance in their parameters across sample characteristics. Unlike CFA's treatment of raw scores as linear measurements, the Rasch model transforms them into logits, offering a more refined assessment (Wright, 1996; Finbråten et al., 2018; Baghaei et al., 2014).

Next, the study utilized Rasch models to validate the Persian adaptation of the Attitudes Toward Online Teaching scale; specifically, the Rasch rating scale model formulated by Andrich (1978) was employed to scrutinize the scale. This meticulous analysis, conducted using the Winsteps® Rasch measurement program and adhering to Linacre's (2017a) guidelines, facilitated a nuanced understanding of participant responses. Besides, Rasch analysis to assess its compliance with the model. Besides, all iterations of the Item Response Theory (IRT) are built upon three foundational assumptions (Baghaei & Christensen, 2023; Meyer, 2014). Monotonicity, the first assumption, dictates that the probability of success or the endorsement of a higher category increases proportionally with the ascending levels of the intended trait (Baghaei & Effatpanah, 2022).

The second assumption, local independence, postulates that respondents' item responses to a scale's items maintain independence at a specific echelon of the intended trait (Forthmann et al., 2020). The third assumption, unidimensionality, mandates that each item within a scale singularly measures a latent trait. Importantly, unidimensionality asserts that the entirety of variability in the data is explicable by a latent trait, establishing a prerequisite for measurement robustness (Stout, 1987). The utility of IRT models extends to facilitating optimal scale development, providing standard errors of measurement for item difficulties and diverse levels of ability/trait among respondents, and evaluating individual items (Doebler & Holling, 2016).

Particularly within this context, exhaustive analysis utilized the unidimensional Rasch rating scale model (Andrich, 1978, 2016; Bradley & Massof, 2017) through Winsteps® Rasch software version 3.73. The dataset comprised 29 items, each assessed on a 5-point ordered response scale: 1 represented 'strongly disagree,' 2 'disagree,' 3 'somewhat agree,' 4 'agree,' and 5 'strongly agree.' The examination meticulously probed a 5-point rating scale from "strongly disagree" to "strongly agree," offering detailed insights into participants' responses and their alignment with the Rasch model. For the purpose of this study, demographic information of the respondents was excluded.

Individual item characteristics and unidimensionality

In the second step of the analysis outlined in Table 1, crucial parameters, including the Infit and Outfit mean-square values, are thoroughly scrutinized to assess the alignment of each of the 29 items with the Rasch model. This detailed scrutiny delves into the difficulty levels, measurement precision, point measure correlations, and adherence to the Rasch model for each individual item.

Here, in Table 1, the "Measure" column elaborates on the difficulty level associated with each item. Higher values indicate increased difficulty for respondents in agreeing with the item's statement. Additionally, the "MODEL S.E." (Standard Error) column elucidates the accuracy of these difficulty measurements. More importantly, smaller SE values indicate more precise estimations, significantly contributing to the overall heightened reliability of the scale. "Point-measure correlations" for all items were additionally calculated to assess the concordance between observed scores and the latent trait. Point-biserial correlations, also referred to as point-

measure correlations, indicate the degree to which responses to each item on the scale align with the overall measure. The results of point-measure correlations, showcased in the table, reveal that all correlations exceed the threshold of 0.30, as stipulated by Linacre (2002).

When evaluating the quality of a measurement tool, ensuring the conformity of its items to the Rasch model is imperative. Analyzing the alignment of item difficulty across a spectrum ensures varying difficulty levels, as per the Rasch model's principle. Examination of response patterns and fit statistics, such as MNSQ Item "Outfit" and MNSQ Item "Infit," helps identify items that might deviate from this anticipated pattern. This rigorous assessment aids in refining the measurement tool, ensuring its validity and accuracy. This principle ideally applies universally to all respondents, irrespective of their skill levels; deviations from this model might suggest that items measure multiple characteristics simultaneously, indicating multi-dimensionality in measurement. Therefore, a thorough examination is essential to uphold the integrity and accuracy of measurements, ensuring alignment with the intended uni-dimensional construct.

Indeed, the identification and rectification of items lacking a clear fit are crucial in measurement analysis. When an item fails to demonstrate a clear fit, two strategies are recommended, i.e., removing the item entirely or replacing it with a new item that aligns more cohesively with the intended measurement construct. Implementing these measures ensures the integrity and validity of the assessment, facilitating an accurate evaluation of the variable under investigation.

Then, assessing item integrity through the examination of 'Infit' and 'Outfit' mean-square (MNSQ) values critically evaluates how effectively items capture the intended construct related to attitudes toward online teaching. The optimal range for these values, recommended by Bond and Fox (2007), is 0.60 to 1.40, with a target value of one. Values above 1.40 may indicate unconventional response patterns, while those below 0.60 might suggest information redundancy, potentially challenging the overall construct validity of the measurement instrument. Elevated values of the infit mean-square (MNSQ) indicate that items are underperforming for the targeted respondents. In fact, this poses a substantial threat to validity, and diagnosing this issue is more challenging compared to addressing high outfit. As elucidated by Linacre (2009b), these two statistical measures convey the extent to which items within the scale accurately reflect the singular latent trait intended for measurement. Misfitting items, not aligned with the construct measured by other items, should be deleted as they introduce construct-irrelevant variance to the data (Baghaei, 2008; Pishghadam et al., 2012).

Table 1 Item Measures and Fit Statistics

ENTRY NUMBER	TOTAL SCORE	TOTAL COUNT	MEASURE	MODEL S.E.	INFIT MNSQ	INFIT ZSTD	OUTFIT MNSQ	OUTFIT ZSTD	PT-MEASURE CORR.	PT-MEASURE EXP.	EXACT OBS%	MATCH EXP%	ITEM
1	592	160	-0.68	0.08	1.09	0.9	1.00	0.0	0.51	0.56	39.4	38.9	I1
2	546	160	-0.37	0.08	1.10	1.0	1.06	0.6	0.48	0.59	44.4	37.5	I2
3	382	160	0.73	0.09	2.75	9.9	2.63	9.9	-0.04	0.62	16.9	38.2	I3
4	488	160	0.00	0.08	1.02	0.2	0.98	-0.1	0.64	0.62	36.9	37.2	I4
5	440	160	0.32	0.08	1.23	2.1	1.14	1.3	0.60	0.63	33.1	37.2	I5
6	524	160	-0.23	0.08	1.96	-0.4	0.88	-1.1	0.67	0.60	43.8	37.0	I6
7	443	160	0.30	0.08	1.14	1.3	1.16	1.4	0.72	0.63	24.4	37.2	I7
8	505	160	-0.11	0.08	1.61	5.2	1.59	4.8	0.42	0.61	41.9	36.2	I8
9	403	160	0.58	0.08	1.44	3.7	1.45	3.6	0.56	0.62	29.4	37.3	I9
10	462	160	0.17	0.08	1.06	0.6	1.02	0.3	0.73	0.62	24.4	37.0	I10
11	533	160	-0.29	0.08	0.81	-2.0	0.79	-2.1	0.67	0.60	37.5	36.6	I11
12	523	160	-0.22	0.08	0.53	-5.5	0.52	-5.4	0.82	0.60	55.6	37.0	I12
13	553	160	-0.42	0.08	0.46	-6.7	0.45	-6.4	0.80	0.59	57.5	38.0	I13
14	472	160	0.11	0.08	0.67	-3.7	0.64	-3.9	0.65	0.62	48.8	36.9	I14
15	336	160	1.09	0.09	0.84	-1.5	0.82	-1.5	0.73	0.60	38.1	41.3	I15
16	440	160	0.32	0.08	1.30	2.6	1.37	3.1	0.35	0.63	35.0	37.2	I16
17	603	160	-0.76	0.08	0.89	-1.0	0.88	-1.0	0.54	0.55	48.8	39.1	I17
18	503	160	-0.09	0.08	0.71	-3.1	0.71	-3.0	0.69	0.61	46.3	36.3	I18
19	490	160	-0.01	0.08	0.60	-4.6	0.59	-4.5	0.76	0.62	48.8	37.1	I19
20	419	160	0.46	0.08	1.60	4.9	1.84	6.3	0.60	0.63	26.3	36.6	I20
21	521	160	-0.21	0.08	1.30	2.8	1.24	2.1	0.59	0.60	39.4	37.0	I21
22	444	160	0.29	0.08	0.66	-3.8	0.65	-3.8	0.86	0.63	40.6	37.2	I22
23	468	160	0.13	0.08	0.46	-6.7	0.47	-6.3	0.71	0.62	60.6	37.0	I23
24	538	160	-0.32	0.08	0.83	-1.8	0.93	-0.6	0.63	0.60	33.8	37.3	I24
25	600	160	-0.74	0.08	0.83	-1.7	0.89	-1.0	0.54	0.55	44.4	39.1	I25
26	539	160	-0.33	0.08	0.85	-1.5	0.92	-0.8	0.54	0.59	40.6	37.3	I26
27	535	160	-0.30	0.08	0.93	-0.7	0.89	-1.0	0.74	0.60	28.8	36.6	I27
28	369	160	0.82	0.09	1.45	3.6	1.47	3.7	0.35	0.61	18.8	38.4	I28
29	527	160	-0.25	0.08	0.74	-2.8	0.69	-3.2	0.69	0.60	43.8	36.5	I29
MEAN	489.6	160.0	0.00	0.08	1.03	-0.3	1.02	-0.3			38.9	37.5	
S.D.	66.8	0.0	0.45	0.00	0.45	3.7	.45	3.7			10.7	1.1	

In the Rasch analysis delineated in Table 1, it becomes apparent that specific items, namely Items 3, 8, 9, 20, and 28, exhibit misfitting characteristics. These items manifest a lack of alignment with the underlying construct measured by other items, as evidenced by their infit and outfit values surpassing the acceptable range. As underscored by Baghaei (2008) and Saghafi et al. (2021), such discrepancies classify these items as poor fits within the scale. Therefore, their exclusion from the scale is imperative to augment its overall adherence to the Rasch model, thereby ensuring its validity and reliability. Therefore, in this study, the identification and subsequent removal of these misfitting items contribute significantly to establishing the unidimensionality of the scale. The adequate data-model fit implies that all items of the scale collectively measure a single latent trait. This observation reinforces the scale's unidimensional nature, signifying its proficiency in effectively capturing the intended trait or attitude under examination.

Item-person map

In the next step of the analysis, a notable aspect of the Rasch model lies in its capacity to express both individual ability and item difficulty estimates on a unified scale, thereby facilitating direct comparison. The relationships between items and individuals are visually represented on an item-person map, known as the Wright map. This map organizes item difficulties and individuals' abilities along a single latent trait line calibrated in logits. This graphical representation effectively highlights the relative difficulty (endorsability) of item estimations on the right side, juxtaposed with the distribution of individuals on the left side. Items positioned at the top of the scale are deemed more challenging to agree with, while those at the bottom are regarded as easier to agree with. Consequently, individuals positioned at the top are perceived to have higher levels of positive attitudes toward online teaching, whereas those at the bottom are viewed as having lower levels of positive attitudes.

However, despite the scale covering a broad range of abilities, there is a noticeable gap in the upper end of the scale. To address this gap effectively, the development of harder items is warranted. Specifically, there is a need for more difficult items with lower agreeability levels to adequately cover the higher end of the scale, where individuals with more positive attitudes toward online teaching are situated. The person-item map further demonstrates that individuals predominantly cluster towards the center of the scale. This clustering suggests that those with

moderate levels of positive attitudes toward online teaching are likely to respond positively to most items in the questionnaire.

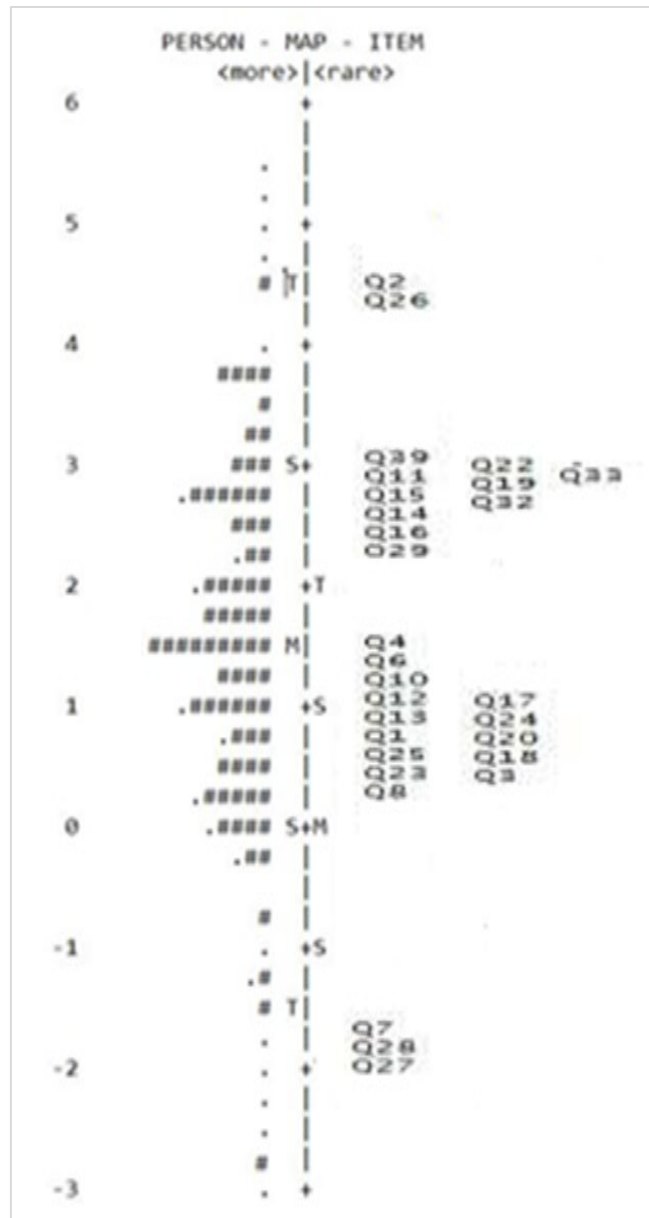


Figure 1 Wright Map of the Distributions of Items and Persons on the latent Trait

Response scale analyses

In the third step of the detailed analysis, a crucial aspect involved scrutinizing participant response distributions across categories, examining the distribution among diverse viewpoints. Category 1, representing the lowest perspective, accounted for approximately 13% of the responses. The "Observed Average (AVRGE)" column reflected mean scores within these categories; that is, Category 1 exhibited an average score of around -0.77. Essential to this assessment was the "Expected Average (EXPECT)," rooted in the Rasch model, which played a crucial role in evaluating participants' anticipated scores in each category.

The evaluation of Rasch model fit highlighted the significance of the 'Infit and Outfit MNSQ' columns, emphasizing their role in assessing how well the model aligns with the data. A fitting alignment, indicating model adherence, typically shows MNSQ values close to 1, with values

surpassing 1.50 raising concerns (Linacre, 2009a). In this instance, all categories comfortably fell within acceptable limits, indicating alignment. The "Andrich Threshold" column revealed specific scale points where participants equally favored one category over another, denoting transition points between viewpoints. Category 1, for instance, had a threshold recorded at -1.26. Lastly, the "Category Measure" column elucidated the placement of each category on the latent trait measured by the scale, with Category 1 assigned a measure of -2.55, indicating the position of each viewpoint on the measured trait. This rigorous scrutiny provided invaluable insights into participant responses within the 5-point rating scale, significantly enhancing our understanding of its evaluative attributes.

Evaluating the quality of the 5-point rating scale involved an in-depth examination of "threshold" values, which are specific points where respondents equally favored adjacent categories. Threshold estimates indicate the level of difficulty associated with distinguishing between each category. The order of the thresholds for items is a crucial factor when evaluating a rating scale. Typically, an increase in threshold estimates with increasing category values is expected, and ordered thresholds play a pivotal role. This expectation stems from the fact that disordered thresholds suggest a poorly defined category for respondents (Linacre, 1999), making it challenging for participants to differentiate between categories.

Next, the analysis presented a systematic sequence of threshold parameters from 1 to 5, signifying "strongly disagree" to "strongly agree," showcasing participants' effective discernment among categories. The anticipated order of the thresholds increasing with category values aligns with established expectations (Bond & Fox, 2007; Linacre, 1999). Andrich threshold values (i.e., -1.26, -0.11, 0.32, and 1.05) affirm the accuracy of participants in distinguishing these respective categories (Table 2); hence, there is no need to break down rating scale categories. As per Table 2, Categories 2, 3, and 4 dominated responses, indicating a moderate inclination toward online teaching and portraying prevailing viewpoints among participants. Fit statistics unequivocally highlighted the scale's suitability in assessing attitudes within this context, demonstrating robust compatibility with the Rasch model's prerequisites of invariance, scalability, and congruity with the data.

Table 2 Category statistics

Category Label	Score	Observed Count	%	Obsvd Avrge	Sample Expect	Infit MnSq	Outfit MnSq	Andrich Threshold	Category Measure	
1	1	586	13	-0.77	-0.74	1.02	1.06	NONE	-2.55	1
2	2	1162	25	-0.33	-0.39	1.02	0.96	-1.26	-0.94	2
3	3	1074	23	-0.03	0.03	0.97	0.96	-0.11	0.05	3
4	4	1024	22	0.55	0.52	0.87	0.81	0.32	0.97	4
5	5	794	17	1.04	1.05	1.14	1.21	1.05	2.42	5

Category probability curve

In the fourth step of the analysis, the visual illustration of the probability curves for each response category was also examined. Figure 2 offers a visual representation of the probability curves for each category, indicating that all categories are anticipated to form a plot resembling a range of hills (Linacre, 2009). In other words, as respondents' levels of the latent variable vary, each category is expected to become more probable than any other category. As depicted in Figure 2, all categories demonstrate proper functionality with distinct peaks. This signifies that each category identifies specific regions of the dimension where respondents are most likely to choose certain response options.

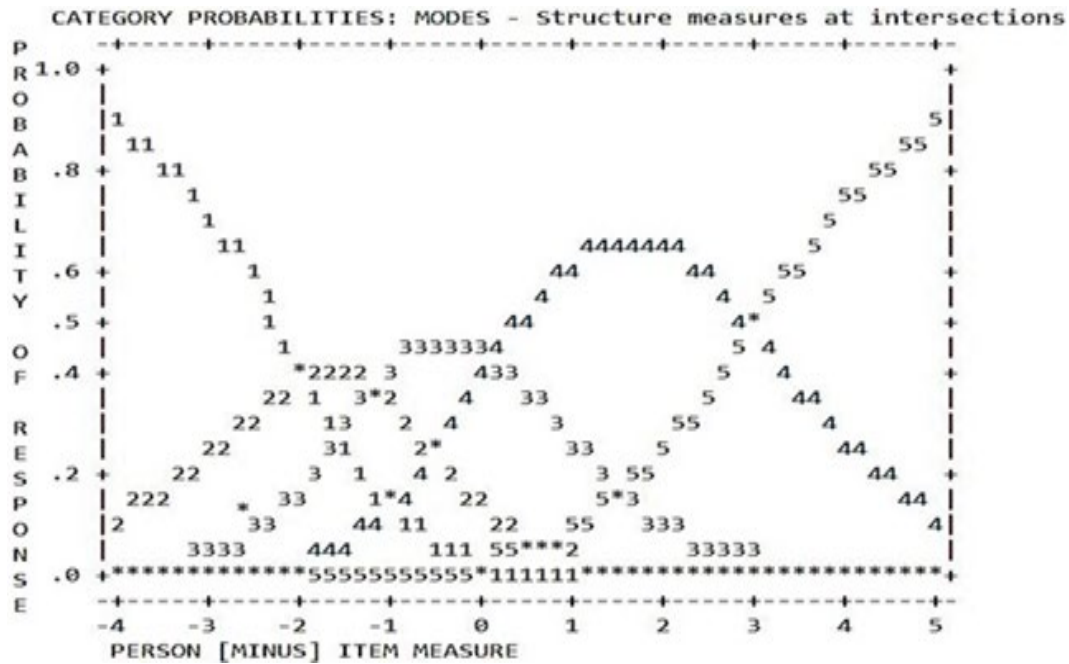


Figure 2 Category probability curves for the scale

Follow-up analysis

In the final step of this subsequent analysis, Items 3, 8, 9, 20, and 28 were excluded; that is, the scale was reanalyzed. The findings from this analysis revealed that the remaining items demonstrated satisfactory outfit and infit values (Sazegar et al., 2018). Alternatively, one may choose to conduct a multidimensional Rasch analysis to determine whether the misfitting items contribute to a distinct and pertinent dimension of the construct (Baghaei, 2012; Baghaei & Aryadoust, 2015), i.e., exploring the possibility of an additional dimension. This analytical approach offers a more nuanced understanding of how the misfitting items relate to the overall construct.

Apart from that, in Rasch analysis, the person separation index takes precedence over traditional reliability indices; that is, a high separation reliability index indicates a strong interrelationship among the test items. Therefore, understanding the overall reliability of the scale is indispensable. The noteworthy Cronbach's alpha reliability score of 0.94 underscores a high level of consistency among items, affirming the scale's dependability. Additionally, the Rasch separation reliability score of 0.92 further attests to the precision and reliability of the scale in capturing the attitudes of Iranian English as a Foreign Language (EFL) teachers toward online teaching.

In closing, following a comprehensive analysis, the findings robustly validate the efficacy of the "Persian Adaptation of the 'Attitudes Toward Online Teaching' Scale," comprising 24 finely-tuned items. Accordingly, this crafted instrument serves as a valuable means for evaluating the viewpoints of Iranian English as a Foreign Language (EFL) teachers regarding online teaching; that is, excelling as a unidimensional tool within the context of Iranian EFL education (Kianiezhad, 2023).

Discussion and conclusion

This study offers a profound exploration of Iranian English as a Foreign Language (EFL) teachers' perspectives on online teaching, representing a comprehensive approach to understanding

evolving educational methodologies. At its core, the research rigorously validated the "Attitudes Toward Online Teaching" scale within the distinctive context of Iranian EFL education, aiming to confirm the scale's relevance in this setting.

The meticulous translation and validation process followed both forward and backward procedures. Subsequent analyses aimed to establish evidence supporting the scale's unidimensional structure and the effectiveness of its 5-point rating system. Next, the Rasch model analysis revealed a shared latent trait influencing participants' responses, affirming the scale's validity.

Ensuring the reliability and relevance of this measurement tool within the specific cultural and educational context of Iranian EFL educators was paramount. The Persian version underwent rigorous validation, taking into account the nuanced perspectives of Iranian EFL teachers to accurately reflect attitudes toward online teaching in this unique setting.

Analyzing the rating scale, it emerged that categories 2, 3, and 4 garnered the majority of responses, indicating a moderate attitude toward online teaching, i.e., bridging attitudes between strong disagreement and strong agreement. This balanced distribution affirms effective differentiation between categories, validating the scale's applicability. Besides, category probability curves show that all the categories have a peak for a certain portion of the scale and are completely ordered.

However, a detailed examination using the Rasch model highlighted deviations for a subset of items—Items 3, 8, 9, 20, and 28—from the optimal range, i.e., raising concerns about their validity in assessing attitudes in the Iranian EFL context and henceforth, it is recommended to remove these items to enhance precision and validity. Furthermore, evaluation of the item-person map showed that the scale effectively targets the respondents. To sum up, the study's overarching findings conclusively affirm that the questionnaire, comprising 24 items, serves as an effective unidimensional tool for assessing the attitudes of Iranian English as a Foreign Language (EFL) teachers toward online teaching (Kianiezhad, 2023). Finally, despite challenges associated with specific items, the overall reliability of the scale remains notably high. The scale demonstrates exceptional reliability with a Cronbach's alpha coefficient of 0.94 and a Rasch separation reliability of 0.92, establishing its pivotal role as a research tool.

The implications of this study are significant for several reasons. Firstly, the rigorous validation of the "Attitudes Toward Online Teaching" scale within the Iranian EFL context provides researchers and educators with a reliable tool for assessing teachers' perspectives on online education. In fact, this has practical implications for designing and implementing effective online teaching strategies in Iranian EFL classrooms.

Secondly, the identification of specific items that deviate from the optimal range highlights areas for improvement in the measurement tool. By removing these items and enhancing the scale's validity, researchers can obtain more accurate insights into teachers' attitudes toward online teaching, leading to better-informed decision-making in educational practice and policy.

Additionally, the study's findings underscore the importance of considering cultural and contextual factors in educational research. By taking into account the nuanced perspectives of Iranian EFL teachers, the study ensures that the measurement tool accurately reflects attitudes toward online teaching within this specific cultural and educational context. Thus, the study contributes to the advancement of knowledge in the field of Iranian EFL education and online teaching. Its findings have implications for both research and practice, providing valuable insights that can inform the development of tailored and effective online education strategies in the Iranian EFL context.

In conclusion, also, with the identified items removed, the validated scale not only confirms its suitability but also identifies areas for refinement. Moreover, its exceptional reliability solidifies its significance as a research tool for tailored and effective online education strategies in the Iranian EFL context.

While this study significantly advances our understanding of Iranian EFL educators' perspectives on online teaching, acknowledging limitations is crucial, i.e., convenience sampling introduces biases, limiting the generalizability of findings. Future research should employ more diversified and representative sampling methodologies, i.e., such as stratified sampling.

Subsequently, incorporating suggested enhancements, like expanding the scale's validation to encompass diverse languages and cultures, would bolster its reliability and pertinence; specifically, examining differential item functioning among subpopulations could further augment the scale's adaptability.

Overall, this study substantially contributes to our knowledge of Iranian EFL teachers' perspectives on online teaching. Embracing methodological refinements and considering broader research inquiries is vital for ongoing progress within the Iranian EFL context. Incorporating these refinements ensures the relevance and applicability of findings in this specific educational setting.

Disclosure statement

No potential conflict of interest was reported by the author(s).

References

- Abdullah, S. A., Saud, M. S., & Mohd, M. H. (2021). Development of M-Learning Implementation Instrument Based on Competency-Based Education: Rasch Analysis. <http://dx.doi.org/10.6007/IJARBS/v11-i8/10549>
- Aryadoust, V., Mehran, P., & Alizadeh, M. (2016). Validating a computer-assisted language learning attitude instrument used in Iranian EFL context: an evidence-based approach. *Computer Assisted Language Learning*, 29(3), 561-595. <https://doi.org/10.1080/09588221.2014.1000931>
- Akaslan, D., & Law, E. L. (2011, April). Measuring teachers' readiness for e-learning in higher education institutions associated with the subject of electricity in Turkey. In *2011 IEEE Global Engineering Education Conference (EDUCON)* (pp. 481-490). IEEE. <https://doi.org/10.1109/EDUCON.2011.5773180>
- Andrich, D. (1978). A rating formulation for ordered response categories. *Psychometrika*, 43, 561-573. <http://dx.doi.org/10.1007/BF02293814>
- Andrich, D. (1988). Rasch models for measurement (Vol. 68). Sage.
- Andrich, D. (2016). Rasch rating-scale model. In *Handbook of item response theory* (pp. 75-94). Chapman and Hall/CRC.
- Asrijanty, A. (2014). Model Rasch sebagai Kerangka Acuan Penyusunan Alat Ukur. *Jurnal Pendidikan dan Kebudayaan*, 20(1), 109-123. <https://doi.org/10.24832/jpnk.v20i1.130>
- Baghaei, P. (2008). The Rasch model as a construct validation tool.
- Baghaei, P. (2009). Understanding the Rasch model.
- Baghaei, P., & Effatpanah, F. (2022). *Elements of psychometrics (2nd Ed.)*. Sokhan Gostar Publishing.
- Baghaei, P., Yanagida, T., & Heene, M. (2017). Development of a Descriptive Fit Statistic for the Rasch Model. *North American Journal of Psychology*, 19(1).
- Baghaei, P., Shoahosseini, R. (2019). A note on the Rasch model and the instrument-based account of validity. *Rasch Measurement Transactions*, 32, 1705-1708.

- Baghaei, P., Hohensinn, C., & Kubinger, K. D. (2014). Persian adaptation of foreign language reading anxiety scale: A psychometric analysis. *Psychological reports*, 114(2), 315-325. <https://doi.org/10.2466/08.11.PR0.114k16w1>
- Baghaei, P., & Christensen, K. B. (2023). Modeling local item dependence in C-tests with the loglinear Rasch model. *Language Testing*, 02655322231155109.
- Baghaei, P., & Cassady, J. (2014). Validation of the Persian translation of the cognitive test anxiety scale. *Sage open*, 4(4), 2158244014555113. <https://doi.org/10.1177/2158244014555113>
- Bradley, C., & Massof, R. W. (2017). Validating translations of rating scale questionnaires using Rasch analysis. *Ophthalmic Epidemiology*, 24(1), 1-2.
- Bond, T. G., & Fox, C. M. (2013). *Applying the Rasch model: Fundamental measurement in the human sciences*. Psychology Press. <https://doi.org/10.4324/9781315814698>
- Borsboom, D. (2008). Latent variable theory.
- Borsboom, D., Mellenbergh, G. J., & Van Heerden, J. (2004). The concept of validity. *Psychological review*, 111(4), 1061.
- Buhl-Wiggers, J., Kjærgaard, A., & Munk, K. (2023). A scoping review of experimental evidence on face-to-face components of blended learning in higher education. *Studies in Higher Education*, 48(1), 151-173.
- Costley, K. C. (2014). The positive effects of technology on teaching and student learning. *Online submission*.
- Dos Santos, L. M. (2024). Experiences and sense-making processes of online teaching internship: A master of teaching programme. *International Journal of Instruction*, 17(1), 23-42.
- Enayati, T., Modanloo, Y., & Kazemi, F. S. M. (2012). Teachers' attitudes towards the use of technology in education. *Journal of Basic and Applied Scientific Research*, 2(11), 010958-010963.
- Finbråten, H. S., Wilde-Larsson, B., Nordström, G., Pettersen, K. S., Trollvik, A., & Guttersrud, Ø. (2018). Establishing the HLS-Q12 short version of the European Health Literacy Survey Questionnaire: latent trait analyses applying Rasch modelling and confirmatory factor analysis. *BMC health services research*, 18(1), 1-17.
- Forthmann, B., Grotjahn, R., Doebler, P., & Baghaei, P. (2020). A comparison of different item response theory models for scaling speeded C-tests. *Journal of Psychoeducational Assessment*, 38(6), 692-705. <https://doi.org/10.1177/0734282919889262>
- Gardner, D. G., Dukes, R. L., & Disenza, R. (1993). Computer use, self-confidence, and attitudes: A causal analysis. *Computers in human behavior*, 9(4), 427-440. [https://doi.org/10.1016/0747-5632\(93\)90033-0](https://doi.org/10.1016/0747-5632(93)90033-0)
- Ghavifekr, S., & Rosdy, W. A. W. (2015). Teaching and learning with technology: Effectiveness of ICT integration in schools. *International journal of research in education and science*, 1(2), 175-191.
- Hung, M. L. (2016). Teacher readiness for online learning: Scale development and teacher perceptions. *Computers & Education*, 94, 120-133. <https://doi.org/10.1016/j.compedu.2015.11.012>
- Jovanova-Mitkovska, S. (2010). The need of continuous professional teacher development. *Procedia-Social and Behavioral Sciences*, 2(2), 2921-2926.
- Kianinezhad, N. (2023) *Examining Iranian EFL Teachers' Attitudes Toward Online Teaching* (Unpublished MA Thesis). Islamic Azad University of Mashhad.
- Kianinezhad, N. (2023). Teacher Development in Technology-Enhanced Language Teaching: Book Review. *Journal of Language and Education*, 9(4), 175-179.
- Kianinezhad, N. (2023). Effective methods of teaching foreign languages online: A global view. *TESOL and Technology Studies*, 4(1), 45-59.
- Kianinezhad, N. (2023). Attitudes towards Online Teaching: Gender, Experience, and Age among Iranian English Language Teaching Teachers. *TESOL and Technology Studies*, 4(2), 15-26.
- Kianinezhad, N. (2024). Educational Background and Workplace Context: Shaping Iranian EFL Teachers' Attitudes Towards Online Teaching. *Studies in Humanities and Education*, 5(1), 29-43.
- Krishnakumar, R., & Rajesh, K. M. (2011). Attitude of Teachers' of Higher Education towards E-learning. *Journal of Education and Practice*, 2(4).

- Kazemi, S., Ashraf, H., Motallebzadeh, K., & Zeraatpishe, M. (2020). Development and validation of a null curriculum questionnaire focusing on 21st century skills using the Rasch model. *Cogent Education*, 7(1), 1736849. <https://doi.org/10.1080/2331186X.2020.1736849>
- Kebritchi, M., Lipschuetz, A., & Santiago, L. (2017). Issues and challenges for teaching successful online courses in higher education: A literature review. *Journal of Educational Technology Systems*, 46(1), 4-29. <https://doi.org/10.1177/0047239516661713>
- Khajavy, G. H., Ghonsooly, B., & Hosseini Fatemi, A. (2018). Development and Validation of the Persian Scale of Teachers' Emotions for the Iranian EFL Teachers. *Language and Translation Studies (LTS)*, 50(4), 73-89. <https://doi.org/10.22067/lts.v50i4.50262>
- Khodabakhshzadehs, H., & Shoahosseini, R. (2021). The Relationships among Attitudes towards Cheating, Academic Self-Confidence, and General Language Ability among Iranian EFL Learners. *International Journal of Language Testing*, 11(2), 1-12.
- Khukalenko, I. S., Kaplan-Rakowski, R., An, Y., & Iushina, V. D. (2022). Teachers' perceptions of using virtual reality technology in classrooms: A large-scale survey. *Education and Information Technologies*, 27(8), 11591-11613. <https://doi.org/10.1007/s10639-022-11061-0>
- Lambri, A., Mahamod, Z., Zakaria, N., Baharum, H., & Hanapi, Z. (2019). Validity and Reliability of the Malay Social Traits Instrument Using Rasch Measurement Model. *International Journal of Academic Research in Business and Social Sciences*, 9(7), 225-232. <http://dx.doi.org/10.6007/IJARBS/v9-i7/6111>
- Lateef, A. R., & Alaba, S. O. (2013). Influence of gender and attitude of pre-service teachers towards on-line instruction in a selected university in South-Western Nigeria. *Asian Social Science*, 9(4), 84. <http://dx.doi.org/10.5539/ass.v9n4p84>
- Linacre, J. M. (2002). What do infit and outfit, mean-square and standardized mean? *Rasch Measurement Transactions*, 16(2), 1-99. <https://www.rasch.org/rmt/rmt162f.htm>
- Linacre, J. M. (2009b). *A user's guide to WINSTEPS*. Chicago, IL: Winsteps.
- Linacre J. M., (2017b), A user's guide to Facets: Rasch-model computer programs, retrieved from <http://www.winsteps.com/manuals.htm>
- Linacre, J. M. (1994). Sample size and item calibration stability. *Rasch measurement transactions*, 7, 328.
- Linacre, J. M. (1999). Investigating rating scale category utility. *Journal of outcome measurement*, 3(2), 103-122.
- Martin, F., Budhrani, K., & Wang, C. (2019). Examining faculty perception of their readiness to teach online. *Online Learning*, 23(3), 97-119. <http://dx.doi.org/10.24059/olj.v23i3.1555>
- Meyer, J. P. (2014). *Applied measurement with jMetrik*. Routledge.
- Moradi, E. (2020). Validation of Persian translation of the intercultural sensitivity scale using Rasch model. *Journal of Ethnic and Cultural Studies*, 7(3), 63-75. <https://doi.org/10.29333/ejecs/374>
- Nadri, M., Baghaei, P., & Zohoorian, Z. (2019). Analysis of the Ruff 2 & 7 Test of Attention with the Rasch Poisson Counts Model. *The Open Psychology Journal*, 12(1).
- Nair, I., & Das, V. M. (2012). Using technology acceptance model to assess teachers' attitude towards use of technology as teaching tool: A SEM approach. *International Journal of Computer Applications*, 42(2), 1-6.
- Ng, D. T. K., Leung, J. K. L., Su, J., Ng, R. C. W., & Chu, S. K. W. (2023). Teachers' AI digital competencies and twenty-first century skills in the post-pandemic world. *Educational Technology Research And Development*, 71(1), 137-161.
- Ningsih, S. K., Mulyono, H., Ar Rahmah, R., & Fitriani, N. A. (2021). A Rasch-based validation of EFL teachers' received online social support scale. *Cogent Education*, 8(1), 1957529. <http://dx.doi.org/10.1080/2331186X.2021.1957529>
- Pishghadam, R., Baghaei, P., & Shayesteh, S. (2012). Construction and validation of an English language teacher creativity scale (ELT-CS). *Journal of American Science*, 8(3), 497-508.
- Papp, R. (1998). Student Perception & Knowledge about Information Technology: A " Computer Attitude and Experience Survey" To Measure Changes. *Journal of Education for MIS*, 5(1), 54-62.
- Petty, R. E., & Krosnick, J. A. (1995). Antecedents and consequences.

- Philip, G., Altbach, L., Reisberg, L., & Rumbley, E. (2009). Trends in global higher education: Tracking an academic revolution. In a report prepared for the UNESCO 2009 world conference on higher education.
- Raja, R., & Nagasubramani, P. C. (2018). Impact of modern technology in education. *Journal of Applied and Advanced Research*, 3(1), 33-35. <http://dx.doi.org/10.21839/jaar.2018.v3iS1.165>
- Ranger, J., & Kuhn, J. T. (2014). Testing fit of latent trait models for responses and response times in tests. *Psychological Test and Assessment Modeling*, 56(4), 382.
- Rasch, G. (1980). (1960/1980). Probabilistic models for some intelligence and attainment tests. (Copenhagen, Danish Institute for Educational Research).
- Stout, W. (1987). A nonparametric approach for assessing latent trait unidimensionality. *Psychometrika*, 52(4), 589-617.
- Sarabi, S. Z. (2017). Validation of the Persian Translation of the Fear of Negative Evaluation Scale. *North American Journal of Psychology*, 19(3), 541-541.
- Sahoo, S. (2020). E-readiness and perception of student teachers' towards online learning in the midst of COVID-19 pandemic. Available at SSRN 3666914. <https://dx.doi.org/10.2139/ssrn.3666914>
- Saghafi, K., Amirian, S. M. R., & Shirvan, M. E. (2021). Differential Item Functioning Analysis of Persian Adaptation of Foreign Language Classroom Anxiety Scale Against Gender. *Human Arenas*, 1-15.
- Salmon, G. (2012). *E-moderating: The key to online teaching and learning*. Routledge.
- Sazegar, Z., Ashraf, H., & Motallebzadeh, K. (2018). Validation of an EFL learner's self-efficacy questionnaire using the Rasch model. *Amazonia Investiga*, 7(16), 23-36.
- Sang, G., Valcke, M., Van Braak, J., & Tondeur, J. (2010). Student teachers' thinking processes and ICT integration: Predictors of prospective teaching behaviors with educational technology. *Computers & Education*, 54(1), 103-112.
- Shirvana, M. E., Khajavya, G. H., & Taherianb, T. (2016). Psychometric analysis of Persian adaptation of foreign language classroom anxiety scale. *English as an Additional*, 195.
- Schwarz, N. (2007). Attitude construction: Evaluation in context. *Social cognition*, 25(5), 638-656.
- Singh, V., & Thurman, A. (2019). How many ways can we define online learning? A systematic literature review of definitions of online learning (1988-2018). *American Journal of Distance Education*, 33(4), 289-306.
- Stickler, U., Hampel, R., & Emke, M. (2020). A developmental framework for online language teaching skills. *Australian Journal of Applied Linguistics*, 3(1), 133-151.
- Sungur Gül, K., & Ateş, H. (2023). An examination of the effect of technology-based STEM education training in the framework of technology acceptance model. *Education and Information Technologies*, 28(7), 8761-8787.
- Suri, G., & Sharma, S. (2016). Investigation of Teacher's attitude towards e-learning-a case study of Panjab University, Chandigarh, India. *Gian Jyoti E-Journal*, 6(3), 1-10.
- Tabatabaee-Yazdi, M., & Baghaei, P. (2022). Persian Translation and Rasch Model-Based Validation of an Intercultural Intelligence Scale. *International Journal of Society, Culture & Language*, 10(1), 71-82.
- Trivedi, S. K., & Yadav, M. (2018). Predicting online repurchase intentions with e-satisfaction as mediator: a study on Gen Y. *VINE Journal of Information and Knowledge Management Systems*, 48(3), 427-447.
- Uzunboylu, H. (2007). Teacher attitudes toward online education following an online inservice program. *International Journal on E-learning*, 6(2), 267-277.
- Wasserman, E., & Migdal, R. (2019). Professional Development: Teachers' Attitudes in Online and Traditional Training Courses. *Online Learning*, 23(1), 132-143.
- Wright, B. D. (1996). Comparing Rasch measurement and factor analysis. *Structural Equation Modeling: A Multidisciplinary Journal*, 3(1), 3-24.
- Yosintha, R., & Yunianti, S. S. (2021). Learner autonomy in EFL online classes in Indonesia: Students' voices. *Langkawi: Journal of The Association for Arabic and English*, 7(1), 119-133.