Iranian Evolutionary Educational Psychology Journal
IEEPJ
University of Hormozgan
Online ISSN: 2588 - 4395
Homepage: https://ieepj.hormozgan.ac.ir

Psychometric Properties of the Teacher Empowering Structure Scale: Classical Test Theory and the Graded Response Model

Masoumeh Bahamin[1], Nooshin Taghinezhad[2]✉, Abdulhadi Samavi[3], Kobra Hajializadeh[4]

1. Ph.D. Student, Department of Psychology, BA.C., Islamic Azad University, Bandar Abbas, Iran
2. Associate Professor, Department of Psychology, BA.C., Islamic Azad University, Bandar Abbas, Iran, nooshintaghinezhad@iau.ac.ir
3. Assistance Professor, Department of Psychology, BA.C., Islamic Azad University, Bandar Abbas, Iran
4. Assistance Professor, Department of Psychology, BA.C., Islamic Azad University, Bandar Abbas, Iran

Article Info

Article type:
Research Article

Article history:
Received 07 Sep. 2025
Received in revised form 19 Oct. 2025
Accepted 28 Oct. 2025
Published online 01 Dec. 2026

Keywords:
Psychometric properties, Teacher Empowering Structure, Classical Test Theory (CTT), Graded Response Model (GRM).

ABSTRACT

Objective: The present study aimed to investigate the psychometric properties of the Teacher Empowering Structure Scale among teachers in Bandar Abbas County, utilizing a psychometric development approach.

Methods: An exploratory mixed-methods design was employed across qualitative and quantitative phases. The qualitative population comprised 8 psychology professors selected via snowball sampling to assess content validity using the Waltz and Bausell method. The quantitative population consisted of teachers in Bandar Abbas County; 500 were selected via multi-stage cluster sampling, with final analyses conducted on 494 individuals. Data were collected using the back-translated Teacher Empowering Structure Scale (Altinkurt & Iliman, 2017). Item analysis was conducted using Classical Test Theory (CTT) and the Graded Response Model (GRM).

Results: Content validity indices indicated satisfactory Content Validity Ratio (CVR = 0.972) and Content Validity Index (CVI) values. Confirmatory Factor Analysis supported a five-factor structure: Participative Decision-Making Environment (9 items), Responsive Environment (4 items), Professional Development Supportive Environment (4 items), Facilitating Environment (7 items), and Autonomy Supportive Environment (6 items). Item factor loadings ranged from +0.41 to +0.93. Based on the GRM, the items exhibited optimal information around the average theta level. Internal consistency analysis showed an overall Cronbach's alpha of 0.876, while IRT marginal reliability coefficients for the subscales ranged from 0.77 to 0.90.

Conclusions: The findings indicate that the Teacher Empowering Structure Scale demonstrates strong validity, reliability, and measurement precision, making it a suitable and robust tool for assessing the empowering structure among teachers in the Iranian population.

Cite this article: Bahamin, M., Taghinezhad, N., Samavi, A., & Hajializadeh, K. (2026). Psychometric properties of the teacher empowering structure scale: Classical test theory and the graded response model. *Iranian Evolutionary Educational Psychology Journal*, 8 (4), 1-27.
DOI: https//doi.org/ 10.22034/8.4.1

Publisher: University of Hormozgan.

Introduction

By designing and implementing high-quality learning processes in educational institutions, teachers play a critical role in enhancing student achievement and facilitating conceptual learning. Teachers constitute a key element in most educational reforms (Fandiño, 2010). Consequently, the quality of a school is fundamentally contingent upon the quality of the teachers employed within it. Teacher empowerment is closely intertwined with the leadership styles of school principals and the opportunities they provide for participative decision-making (Eddy, 2009). This is primarily because school principals serve as the core facilitators of both student and teacher empowerment within the school environment (Morales, 2015).

The concept of teacher empowerment has been explored through various definitions by different researchers. According to Sharma (2014), empowering teachers entails supporting them to act as active shapers by validating their experiences, decision-making authority, and power, thereby fostering a sense of being true key figures in school practices and conditions. Bogler and Somech (2004) define teacher empowerment as a process in which teachers are actively engaged in their own professional growth and possess the capacity to resolve their own challenges. Essentially, teacher empowerment involves the self-regulation of one's personality, cognition, and motivation. Furthermore, Zimmerman (2000) argues that teacher empowerment serves as both a process and an outcome.

Empirical research findings have demonstrated that teacher empowerment generally exerts a positive influence within educational settings. For instance, scholars have found that teacher empowerment enhances teachers' job satisfaction (Rice & Schneider, 1994; Rinehart & Short, 1994), professional commitment, and organizational citizenship behaviors (Bogler & Somech, 2004), as well as organizational commitment (Somech, 2005). It also fosters professionalism and self-confidence while concurrently mitigating teacher burnout (Day et al., 2003). Therefore, it is postulated that empowering teachers and awakening their sense of efficacy can elicit numerous positive organizational behaviors, ultimately playing a pivotal role in organizational success and the career sustainability of teachers (Bogler & Somech, 2004).

The concept of teacher empowerment has been explored across various dimensions by different researchers. Wilson and Coolican (1996) conceptualize teacher empowerment in two dimensions: external and internal power. Short and Rinehart (1992) discuss teacher empowerment across six

dimensions: "Decision Making," "Professional Growth," "Status," "Self-Efficacy," "Autonomy," and "Impact." Yin et al. (2009) consider teacher empowerment in three dimensions: "Professional growth in school," "Participation in decision-making," and "The impact of teachers' work on other colleagues." Al-Yaseen and Al-Musallam (2015) reviewed the literature on teacher empowerment and identified 13 dimensions: Accountability, Authority, Curriculum planning, Collaboration, Decision making, Impact, Professional development, Professional knowledge, Responsibility, Self-efficacy, Self-esteem, Status, and New teacher training.

Altinkurt et al. (2016) state that the concept of teacher empowerment comprises two main dimensions: *structural empowerment*, which focuses on administrative processes and their regulation, and *psychological empowerment*, which drives teachers' perceptions. In the international literature, it is evident that Kanter's (1993) scale is frequently used to assess structural empowerment, while Spreitzer's (1995) scale is widely used for psychological empowerment. Kanter (1993) addresses and explains structural empowerment in terms of access to information, opportunity, resources, support, formal power, and informal power. Spreitzer (1995), on the other hand, discusses and elucidates psychological empowerment in terms of meaning, impact, competence, and self-determination (autonomy).

A review of the literature concerning the definitions and classifications of teacher empowerment reveals a lack of consensus. It is observed that the most widely used data collection instrument related to this topic is Short and Rinehart's (1992) Teacher Empowerment Scale. The original name of Short and Rinehart's (1992) instrument is the "School Participant Empowerment Scale" (SPES). While this scale is referred to as the "Teacher Empowerment Scale" in some studies (Okman, 2018; Somech, 2005), it is considered the "School Participant Empowerment Scale" in others (Bogler & Nir, 2012; Jiang et al., 2019). Furthermore, Short and Rinehart's (1992) scale was conducted on an Israeli sample in 1992. The "Teacher Empowerment Scale" developed by Yin et al. (2009) in the context of curriculum reform in China utilized a sample of Chinese teachers.

Regarding structural empowerment, a study by Pelit et al. (2011) was conducted using a different scale; a scale development study for teacher empowerment was carried out by Yeliz and Abdurrahman (2020); and a qualitative study focusing on empowering school environments was conducted by Balkar (2015). Therefore, it can be concluded that studies on structural empowerment are predominantly conducted in the fields of health sciences and business

administration, with educational science studies often drawing upon them. However, it is argued that structural empowerment scales developed for business administration and health sciences are inadequate for assessing the environments of educational organizations, as schools are loosely structured and value-driven organizations. Moreover, educational services in schools are not merely routine tasks that can be accomplished through individual effort and decision-making; rather, these services inherently demand creativity and teamwork. On the other hand, teachers constitute almost the entirety of the school staff and are professionals who have been formally educated in their respective fields (Altinkurt & Yılmaz, 2013). For these reasons, the need to develop a new instrument for assessing teachers' structural empowerment arises.

Beyond all this, the concept of teacher empowerment frequently encountered in the international literature generally deals with both psychological and structural empowerment simultaneously. When studies related to this concept are analyzed (Bogler & Nir, 2012; Bogler & Somech, 2004; Inam, 2015; Kim et al., 2014; Rinehart et al., 1998), a six-dimensional structure is observed. Decision-making, professional growth, status, self-efficacy, autonomy, and impact form the foundational structural skeleton through which teacher empowerment is analyzed. Some of these dimensions pertain to psychological empowerment, while others relate to structural empowerment. Undoubtedly, teacher empowerment can be evaluated holistically by considering both psychological and structural aspects together. However, addressing these concepts independently provides a more substantial contribution to solving the problem. Furthermore, teacher empowerment is fundamentally context-dependent. For this reason, deviating from the broad classification of teacher empowerment, the specific conditions of empowering environments are taken into account, and the scope of Altinkurt and Yılmaz's (2017) research is strictly limited to structural empowerment. In this context, teachers' structural empowerment is conceptualized as a five-factor structure comprising Participative Decision-Making Environment, Responsive Environment, Professional Development Supportive Environment, Facilitating School Environment, and Autonomy Supportive Environment.

According to Altinkurt and Yılmaz (2017), their study aimed to investigate and develop a data collection instrument to define the levels of teachers' structural empowerment. The research sample consisted of elementary, middle, and high school teachers working in Muğla, Turkey, who voluntarily participated in the study. The scale development procedure was implemented on two

independent and distinct samples. Data were collected during the 2015-2016 academic year. Exploratory Factor Analysis (EFA) was conducted using the data collected from the first sample, while Confirmatory Factor Analysis (CFA) was performed on the data obtained from the second sample.

The first sample comprised 261 teachers, of which 33.3% were employed in elementary schools, 31.8% in middle schools, and 34.9% in high schools. The second sample included 331 teachers. Within this group, 24.9% worked in elementary schools, 41.4% in middle schools, and 33.7% in high schools. Factor analysis was conducted to establish construct validity. The five-factor structure that emerged as a result of the EFA included: Participative Decision-Making Environment, Responsive Environment, Professional Development Supportive Environment, Facilitating School Environment, and Autonomy Supportive Environment. This five-factor structure accounted for 59.88% of the total variance.

The scale consists of 30 items rated on a five-point Likert scale ranging from "1-Strongly disagree" to "5-Strongly agree." The structure was further examined through CFA, and the results indicated that the scale demonstrated good model fit. To establish the reliability of the scale, item-total correlations, Cronbach's alpha internal consistency coefficients, and the item means of the upper and lower 27% groups were examined. Consequently, a psychometrically adequate, valid, and reliable data collection instrument was developed to assess the structural empowerment of teachers (Altinkurt & Yılmaz, 2017).

Therefore, addressing the limitations of previous studies—most notably the underutilization of advanced psychometric approaches, such as the Graded Response Model (GRM), for estimating item and overall scale parameters—the current study aims to validate and analyze the psychometric properties of Altinkurt and Yılmaz's (2017) Teacher Empowering Structure Scale within the target population of teachers in Bandar Abbas. Consequently, building upon these identified gaps in the literature, this research seeks to answer the following critical question: Are the psychometric properties and reliability of Altinkurt and Yılmaz's (2017) Teacher Empowering Structure Scale satisfactory among teachers in Bandar Abbas when evaluated through the lenses of the Graded Response Model (GRM) and Classical Test Theory (CTT)?

Material and Methods

Research Design and Participants

The present study employed an exploratory sequential mixed-methods design with a psychometric developmental approach. To this end, an initial qualitative study was conducted wherein the face validity, Content Validity Ratio (CVR), and Content Validity Index (CVI) of Altinkurt and Yılmaz's (2017) Teacher Empowering Structure Scale were evaluated by 8 psychology professors based on the Waltz and Bausell method. Subsequently, in the quantitative phase, the instrument was validated to measure the teacher empowering structure. The psychometric properties of the questionnaire were examined utilizing Classical Test Theory (CTT) and the Graded Response Model (GRM). The GRM is formulated to estimate individuals' true scores in questionnaires featuring graded (polytomous) response categories. This model allows researchers to determine the probability that an examinee possessing a specific level of a latent trait will endorse a particular response option (Nering & Ostini, 2011).

The statistical population for the qualitative phase comprised all psychology professors in Bandar Abbas. A snowball sampling method was utilized to select the sample. Snowball sampling identifies potential participants or key informants by having each psychology expert refer the subsequent potential candidate to the researchers. Inclusion criteria for this phase required accessibility, holding a Ph.D., active engagement in teaching and research at academic institutions, and familiarity with the concept of the teacher empowering structure. In qualitative studies, according to Lawshe's (1975) table, the minimum sample size required to assess face validity, CVR, and CVI is 5 experts; however, to ensure robustness, 8 psychology experts were selected for the present study.

For the quantitative phase, the statistical population included all teachers in Bandar Abbas, selected via a multi-stage cluster sampling method. Inclusion criteria for the quantitative phase required being an active, in-service teacher in Bandar Abbas and demonstrating a willingness to participate in the study. Regarding sample size in psychometric studies, Comrey and Lee (2013) recommend 10 participants per item. Kline (2023) suggests 10 to 20 subjects per variable for Exploratory Factor Analysis (EFA); conversely, the minimum sample size for Confirmatory Factor Analysis (CFA) is determined by the number of factors rather than variables, equating to 10 to 20 subjects per factor. Furthermore, Hooman (2011) estimates that for factor analysis, a sample size of 100 is

poor, 200 is fair, 300 is good, 500 is very good, and 1,000 is excellent. Concerning appropriate sample sizes for Item Response Theory (IRT) models, findings by Jiang et al. (2016) indicated that a sample size of 500 is adequate for accurate parameter estimation. Therefore, accounting for potential participant attrition, 500 individuals were initially selected as the sample. Following data collection and the elimination of incomplete questionnaires, the data of 494 participants were subjected to statistical analysis.

Instruments

Teacher Empowering Structure Scale: The Teacher Empowering Structure Scale was developed and validated by Altinkurt and Yılmaz (2017). This scale comprises 30 items, each rated on a five-point Likert scale ranging from "1 = Strongly disagree" to "5 = Strongly agree." It features a five-factor structure, which includes: Participative Decision-Making Environment (9 items), Responsive Environment (4 items), Professional Development Supportive Environment (4 items), Facilitating School Environment (7 items), and Autonomy Supportive Environment (6 items).

In a study conducted among elementary, middle, and high school teachers in Turkey, Altinkurt and Yılmaz (2017) utilized Exploratory Factor Analysis (EFA) and Confirmatory Factor Analysis (CFA) to assess construct validity. The EFA revealed a five-factor structure—comprising the Participative Decision-Making Environment, Responsive Environment, Professional Development Supportive Environment, Facilitating School Environment, and Autonomy Supportive Environment—which accounted for 59.88% of the total variance. Furthermore, CFA results demonstrated good model fit for the scale. To evaluate the scale's reliability, item-total correlations, Cronbach's alpha internal consistency coefficients, and the item means of the upper and lower 27% groups were examined. Consequently, it was established as a psychometrically valid and reliable data collection instrument for assessing teachers' structural empowerment, with Cronbach's alpha coefficients for internal consistency ranging from 0.77 to 0.95.

Additionally, Akan (2021) investigated the relationship between learning-centered leadership and teachers' structural empowerment using a sample of 261 teachers working in secondary education institutions in the central district of Bolu province, Turkey, during the 2020-2021 academic year. Akan employed Altinkurt and Yılmaz's (2017) scale to measure structural empowerment. In Akan's study, the overall Cronbach's alpha reliability coefficient for the scale was calculated as 0.97. The subscale alphas were reported as 0.94 for the Participative Decision-Making

Environment, 0.83 for the Responsive Environment, 0.82 for the Professional Development Supportive Environment, 0.91 for the Facilitating School Environment, and 0.90 for the Autonomy Supportive Environment.

Psychometric Properties in the Current Study

In the present study, the Teacher Empowering Structure Scale was administered in Iran for the first time. Using the Waltz and Bausell method, the Content Validity Ratio (CVR) for the scale was calculated as 0.972. The Content Validity Index (CVI) was assessed across three criteria, yielding scores of 0.96 for clarity, 0.98 for simplicity, and 0.96 for relevance of all items.

The assumption of local independence was met based on Pearson's chi-square index, applying Samejima's Graded Response Model. Additionally, the one-dimensionality assumption was confirmed through Multidimensional Item Response Theory (MIRT) analysis and Exploratory Factor Analysis using Principal Axis Factoring (PAF).

Based on Classical Test Theory (CTT) and the Graded Response Model (GRM), the item parameters were found to be optimal. The factor loadings for the scale items ranged from +0.41 to +0.93, with item 30 exhibiting the lowest and item 15 exhibiting the highest factor loading. Specifically:

- **Participative Decision-Making Environment:** item 7 had the lowest (+0.60) and item 2 the highest (+0.84) loading.
- **Autonomy Supportive Environment:** item 30 had the lowest (+0.41) and item 25 the highest (+0.83) loading.
- **Responsive Environment:** item 10 had the lowest (+0.86) and item 13 the highest (+0.92) loading.
- **Professional Development Supportive Environment:** item 17 had the lowest (+0.87) and item 15 the highest (+0.93) loading.
- **Facilitating School Environment:** item 21 had the lowest (+0.61) and item 18 the highest (+0.72) loading.

The Test Information Function (TIF) peaked around the average theta. To estimate the reliability of the questionnaire's components using the marginal method based on Item Response Theory, the marginal reliabilities were calculated as 0.87 for the Participative Decision-Making Environment, 0.88 for the Responsive Environment, 0.90 for the Professional Development Supportive

Environment, 0.82 for the Facilitating School Environment, and 0.77 for the Autonomy Supportive Environment.

Furthermore, examining the internal consistency of the 30-item questionnaire based on Classical Test Theory, the overall Cronbach's alpha coefficient was calculated as 0.876. The alpha coefficients for the subscales were calculated as follows: 0.878 for the Participative Decision-Making Environment (9 items), 0.793 for the Autonomy Supportive Environment (6 items), 0.922 for the Responsive Environment (4 items), 0.891 for the Professional Development Supportive Environment (4 items), and 0.812 for the Facilitating School Environment (7 items).

Procedure

To conduct this study, the Teacher Empowering Structure Scale developed by Altinkurt and Yılmaz (2017) was first translated into Persian by an English language expert using a forward-backward translation process. Subsequently, the back-translated version was compared with the original text. Any discrepancies or translation issues concerning the items were reviewed and resolved by translation experts and psychology professors.

The face validity, Content Validity Ratio (CVR), and Content Validity Index (CVI) of the translated scale were evaluated by 8 psychology professors from universities in Bandar Abbas, based on the Waltz and Bausell method. Specifically, the expert panel rated the "simplicity," "clarity," and "relevance" of each item using a 4-point Likert scale. According to Lawshe's (1975) table, as cited in Kiani (2022), the minimum acceptable value for CVR and CVI with a panel of 8 evaluating experts is 0.75.

Furthermore, the translated scale was administered to a preliminary pilot sample of 30 teachers in Bandar Abbas to assess the comprehensibility, clarity, and simplicity of the items. Following data collection from this pilot group, necessary linguistic and structural revisions were implemented. Ultimately, the finalized scale was administered to a primary sample of 500 teachers in Bandar Abbas. After data collection and the elimination of incomplete questionnaires, the psychometric properties of the scale were analyzed using the data from the remaining 494 participants. The statistical analyses were conducted based on Classical Test Theory (CTT) and the Graded Response Model (GRM), utilizing IRT-PRO, SPSS (version 27), NOHARM, and LISREL software packages.

Ethical Considerations

This study was conducted in accordance with the ethical principles of human research and complied with the guidelines of the University Research Ethics Committee. Participation in the study was voluntary, and all participants were informed about the objectives and procedures of the research prior to data collection. Informed consent was obtained from all participants. They were assured of the confidentiality and anonymity of their responses and informed of their right to withdraw from the study at any stage.

Results

In the qualitative phase of the study, 8 psychology professors from universities and higher education institutions in Bandar Abbas participated, while the quantitative phase involved 494 teachers from Bandar Abbas.

In the qualitative phase, the face validity, Content Validity Ratio (CVR), and Content Validity Index (CVI) of the Teacher Empowering Structure Scale were examined. To establish the face validity of the questionnaire, the feedback of the 8 psychology experts was reviewed regarding aspects such as wording, comprehensibility, appropriateness, and the avoidance of technical jargon in the items. Based on their recommendations, the necessary revisions were implemented.

Following the distribution of the scale to the 8 psychology experts to assess the CVR, each item was rated based on a three-point scale ("Not necessary" = 0, "Useful" = 1, "Essential" = 2). The overall CVR for the Teacher Empowering Structure Scale was calculated as 0.972; the detailed results of this analysis are reported in Table 1.

Table 1. Content Validity Results for the Teacher Empowering Structure Scale

Component	Item	Score	CVR
Participative Decision-Making	1	16	1.00
	2	16	1.00
	3	15	0.88
	4	16	1.00
	5	16	1.00
	6	16	1.00
	7	15	0.88
	8	16	1.00
	9	16	1.00
Responsive Environment	10	16	1.00
	11	16	1.00
	12	15	0.88
	13	16	0.88

Professional Development Supportive Environment	14	16	1.00
	15	16	1.00
	16	15	0.88
	17	16	1.00
Facilitating Environment	18	16	1.00
	19	16	1.00
	20	15	0.88
	21	15	0.88
	22	16	1.00
	23	16	1.00
	24	16	1.00
Autonomy Supportive Environment	25	16	1.00
	26	16	1.00
	27	16	1.00
	28	16	1.00
	29	15	0.88
	30	16	1.00

Furthermore, the Content Validity Index (CVI) was calculated for the criteria of clarity, simplicity, and relevance, yielding values of 0.96, 0.98, and 0.96, respectively. The detailed results of this analysis are presented in Table 2.

Table 2. Content Validity Index (CVI) Results for the Teacher Empowering Structure Scale

Component	Item	Clarity	Simplicity	Relevance
Participative Decision-Making	1	1.00	1.00	1.00
	2	1.00	1.00	0.80
	3	0.80	1.00	1.00
	4	1.00	0.80	1.00
	5	1.00	1.00	1.00
	6	1.00	1.00	0.80
	7	1.00	1.00	1.00
	8	0.80	1.00	1.00
	9	1.00	1.00	1.00
Responsive Environment	10	1.00	1.00	1.00
	11	1.00	1.00	0.80
	12	1.00	1.00	1.00
	13	0.80	1.00	1.00
Professional Development Supportive Environment	14	1.00	1.00	1.00
	15	1.00	1.00	1.00
	16	1.00	0.80	1.00
	17	0.80	1.00	1.00
Facilitating Environment	18	1.00	1.00	1.00
	19	1.00	1.00	0.80
	20	1.00	1.00	1.00
	21	0.80	1.00	1.00
	22	1.00	1.00	1.00
	23	1.00	1.00	1.00
Autonomy Supportive Environment	25	1.00	1.00	0.80
	26	0.80	1.00	1.00
	27	1.00	0.80	1.00
	28	1.00	1.00	1.00
	29	1.00	1.00	1.00

Prior to conducting the quantitative analyses, the overall status of the dataset was screened. Initially, out-of-range values were identified, and necessary corrections regarding data entry were made. Subsequently, missing data were evaluated; questionnaires with more than 5% missing responses across items were excluded from the analysis. After verifying the statistical assumptions, the final analyses of the research questions were performed on the remaining sample of 494 participants. First, item analysis was conducted based on Classical Test Theory (CTT). In this section, the psychometric properties of the Teacher Empowering Structure Scale were examined using CTT, and the corresponding results are presented in Table 3.

Table 3. Descriptive Statistics, Slope, and Reliability Coefficients of the Teacher Empowering Structure Scale Items

Component	Item	Mean	*SD*	Slope	Slope/α\alphaα
Participative Decision-Making	1	2.07	1.37	0.658	0.862
	2	1.86	1.35	0.654	0.862
	3	2.29	1.42	0.673	0.860
	4	2.21	1.35	0.662	0.862
	5	2.86	1.38	0.620	0.865
	6	2.94	1.38	0.621	0.865
	7	2.49	1.52	0.565	0.871
	8	2.31	1.35	0.594	0.867
	9	2.78	1.42	0.549	0.871
Responsive Environment	10	2.68	1.27	0.433	0.822
	11	2.49	1.25	0.643	0.718
	12	2.62	1.14	0.645	0.721
	13	2.75	1.27	0.702	0.686
Professional Development Supportive Environment	14	2.68	1.22	0.801	0.905
	15	2.49	1.23	0.832	0.894
	16	2.53	1.29	0.846	0.890
	17	2.64	1.22	0.802	0.905
Facilitating Environment	18	2.72	1.32	0.589	0.780
	19	2.78	1.32	0.578	0.782
	20	2.67	1.33	0.564	0.784
	21	2.73	1.43	0.512	0.793
	22	2.57	1.37	0.539	0.789
	23	2.83	1.29	0.559	0.785
	24	2.53	1.36	0.496	0.796
Autonomy Supportive Environment	25	3.50	1.25	0.574	0.602
	26	3.17	1.26	0.498	0.628
	27	2.87	1.19	0.351	0.674
	28	2.28	1.27	0.431	0.650
	29	3.87	1.29	0.369	0.670
	30	2.87	1.38	0.329	0.685

Based on the results presented in Table 3, the item discrimination indices for the Teacher Empowering Structure Scale, according to Classical Test Theory (CTT), ranged from +0.329 to +0.846. Item 30 exhibited the lowest discrimination index, while Item 16 demonstrated the highest.

Within the Participative Decision-Making Environment component, Item 3 had the highest index (0.673) and Item 7 had the lowest (0.565). For the Autonomy Supportive Environment component, Item 25 had the highest (0.574) and Item 30 had the lowest (0.329). In the Responsive Environment component, Item 10 recorded the highest index (0.702) and Item 10 the lowest (0.433). For the Professional Development Supportive Environment component, Item 16 showed the highest index (0.846) and Item 14 the lowest (0.801).

The internal consistency of the items was also evaluated based on CTT. The internal consistency reliability was computed in two phases. First, all items (16 items) were analysed simultaneously, yielding a Cronbach's alpha coefficient of 0.876 for the total scale. The alpha coefficients for the subscales were as follows: Participative Decision-Making Environment (9 items) = 0.878; Autonomy Supportive Environment (6 items) = 0.793; Responsive Environment (4 items) = 0.922; Professional Development Supportive Environment (4 items) = 0.891; and Facilitating Environment (7 items) = 0.812.

Additionally, the contribution of each item to the internal consistency was examined using the "alpha if item deleted" method. The results indicated that the deletion of any single item did not lead to an increase in the internal consistency of its respective component. Therefore, all items within each dimension are homogeneous and exhibit strong correlations with their corresponding components. Consequently, the items of the Teacher Empowering Structure Scale demonstrated satisfactory reliability based on CTT.

Subsequently, the item analysis based on the Graded Response Model (GRM) is reported. In this section, the psychometric properties of the scale were evaluated using the GRM framework. The GRM is capable of analysing categorical responses along an ordinal continuum, such as a Likert scale, where respondents' tendencies or attitudes are assessed through a single statement with multiple response categories. In this context, subjects' attitudes and opinions are expressed along a continuum via discrete categories. The GRM assumes that an item response is an ordered categorical variable where the scores are not separated by equal intervals (Lee, 2013).

Therefore, given that the response pattern of the Teacher Empowering Structure Scale consists of polytomous categorical items scored on a Likert scale, the GRM was employed to examine its psychometric properties. Furthermore, the IRTPRO software was utilized to assess the

assumptions of local independence and one-dimensionality, as well as to estimate the item information functions for the scale.

Before analysing the items using Item Response Theory (IRT), it is necessary to ensure that two fundamental assumptions of IRT, namely local independence and one-dimensionality, are met. To investigate the assumption of local independence among the items of the Teacher Empowering Structure Scale, the values of Pearson's chi-square and marginal chi-square indices were calculated using IRTPRO software, applying Samejima's Graded Response Model (GRM) and assuming a single dimension for all items of each characteristic. The calculated values for all items were less than 10, indicating that the assumption of local independence among the items was satisfied. Therefore, in the Teacher Empowering Structure Scale, the response to each item is independent, and answering or not answering one item has no influence on the responses to other items.

The other core assumption of IRT is one-dimensionality. To assess the one-dimensionality assumption of the scale, Multidimensional Item Response Theory (MIRT) analysis was utilized. In this method, two models—unidimensional and two-dimensional—were analysed to infer the dimensionality of the test by comparing their fit. Given that the calculated Root Mean Square Residual (RMSR) for the scale in the unidimensional model showed a reduction of less than 10% compared to the two-dimensional model, and considering the calculated Tanaka index for the unidimensional (0.9301474) and two-dimensional (0.9215304) models, it was concluded that the Teacher Empowering Structure Scale is unidimensional.

To further investigate one-dimensionality, Exploratory Factor Analysis (EFA) using Principal Axis Factoring was also employed. In this method, the first factor must explain at least 20% of the variance, and the eigenvalue of the first factor is compared to that of the second factor (Reckase, 2009). To evaluate data suitability for factor analysis, two statistical indices are necessary: the Kaiser-Meyer-Olkin (KMO) measure and Bartlett's Test of Sphericity. The KMO index determines whether the sample size is adequate for factor analysis. The closer this index is to 1, the more adequate the collected data is for analysis. Bartlett's Test of Sphericity is the second test conducted prior to factor analysis to check whether the correlation matrix of the collected data is statistically significant. This significance is evaluated via the chi-square distribution. If the null hypothesis (H_0) in Bartlett's test is rejected, conducting factor analysis is justified. The foundation of

factor analysis is the presence of significant correlations among test items; otherwise, applying factor analysis techniques would be ineffective.

Data analysis for the Teacher Empowering Structure Scale revealed a KMO measure of sampling adequacy of 0.867. Furthermore, Bartlett's Test of Sphericity yielded a value of 6100.834 with 435 degrees of freedom, which was statistically significant at the 99% confidence level. This indicates that the sample size is sufficient for factor analysis and the population correlation matrix is not zero. Based on both criteria, it is justifiable to conduct factor analysis on the studied sample.

Moreover, based on the results of the exploratory factor analysis using principal components, the first extracted factor from the data matrix explained 22.13% of the variance. The second factor explained 11.20%, the third 11.09%, the fourth 8.40%, and the fifth 7.06%. Altogether, these five factors accounted for 59.88% of the total variance. According to Reckase (2009), this satisfies the assumption of one-dimensionality. Based on the report by Kirisci et al. (2001), the application of a unidimensional model is acceptable when there is one dominant factor alongside several minor factors within the data matrix. By summarizing the results of the RMSR analysis and the exploratory factor analysis, it can be stated that the data derived from the Teacher Empowering Structure Scale reflect a unidimensional model.

After confirming the local independence and one-dimensionality of the data matrix, the scale was analysed based on the Graded Response Model using IRTPRO software. The analysis results are presented in Table 4, and the item information functions and category response curves for the components of the Teacher Empowering Structure Scale are displayed in Figure 1.

Table 4. Item Parameters of the Teacher Empowering Structure Scale Based on the Graded Response Model

Component	*i*	*a*	*b*1	*b*2	*b*3	*b*4	λ\lambdaλ
Participative Decision-Making	1	2.48	-0.11	0.57	0.98	1.72	0.83
	2	2.62	-0.39	0.78	1.15	1.63	0.84
	3	2.29	-0.18	0.40	0.86	1.59	0.80
	4	2.03	-0.18	0.50	0.97	1.87	0.77
	5	1.64	-1.01	-0.24	0.52	1.60	0.69
	6	1.73	-1.04	-0.32	0.37	1.55	0.71
	7	1.28	-0.24	0.25	0.80	1.94	0.60
	8	1.63	-0.36	0.36	1.18	1.95	0.69
	9	1.34	-1.01	-0.02	0.58	1.72	0.62
Responsive Environment	10	2.83	-0.88	0.16	1.93	1.93	0.86
	11	3.19	-0.73	0.42	2.01	2.01	0.88
	12	3.52	-1.02	0.28	2.03	2.03	0.90
	13	5.29	0.89	0.28	2.24	2.24	0.92
Professional Development Supportive Environment	14	3.76	-0.99	-0.03	1.31	1.31	0.91

	15	4.33	-0.75	0.24	1.50	1.50	0.93
	16	4.44	-0.67	0.17	1.35	1.35	0.93
	17	2.99	-0.96	-0.01	1.64	1.64	0.87
Facilitating Environment	18	1.74	-0.87	-0.33	1.66	1.66	0.72
	19	1.69	-0.98	-0.26	1.71	1.71	0.70
	20	1.67	-0.90	0.02	1.71	1.71	0.70
	21	1.31	-0.84	-0.24	1.69	1.69	0.61
	22	1.46	-0.74	0.08	1.77	1.77	0.65
	23	1.56	-1.15	-0.39	1.76	1.76	0.68
	24	1.43	-0.64	0.09	1.92	1.92	0.64
Autonomy Supportive Environment	25	2.25	-1.57	-0.96	-0.21	0.88	0.83
	26	1.73	-1.67	-0.69	0.20	1.38	0.71
	27	0.81	-2.28	-0.80	1.13	3.14	0.43
	28	1.04	-2.34	-1.10	0.02	1.63	0.52
	29	1.01	-2.66	-1.99	-1.02	0.33	0.51
	30	0.77	-1.81	-0.42	0.79	2.44	0.41

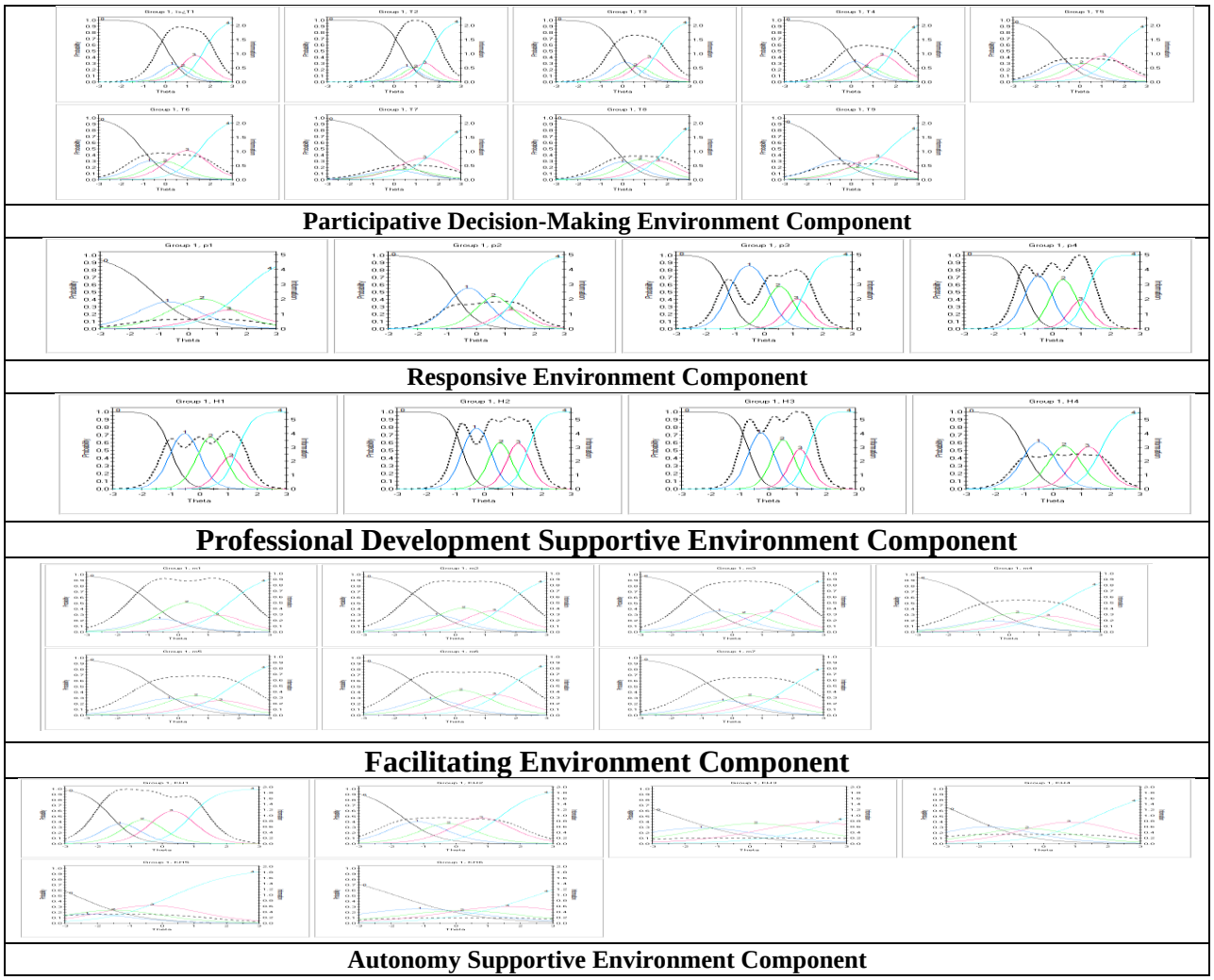

Figure 1. Item Information Functions and Category Response Curves for the Components of the Teacher Empowering Structure Scale

An evaluation of the item fit indices based on the Graded Response Model (GRM) indicated that the calculated chi-square significance level for all items was greater than 0.05; consequently, these items demonstrate adequate fit indices.

The item discrimination parameters for the total scale ranged from +0.77 to +5.29. For the components, these parameters ranged from +1.34 to +2.62 for the Participative Decision-Making Environment; +0.77 to +2.25 for the Autonomy Supportive Environment; +2.83 to +5.29 for the Responsive Environment; +2.99 to +4.44 for the Professional Development Supportive Environment; and +1.31 to +1.74 for the Facilitating Environment.

Furthermore, the first threshold parameter ranged from -2.66 to +0.39; the second threshold parameter ranged from -1.99 to +0.78; the third threshold parameter ranged from -1.02 to +2.24; and the fourth threshold parameter ranged from +0.33 to +2.44. Based on the findings in Table 4 and Figure 1, the Category Response Curves (CRCs) exhibited a well-ordered pattern, such that their distributions are clearly distinguishable from one another. Additionally, the curves are not flat and do not overlap or intertwine unnecessarily.

The thresholds for the response categories of all items are sufficiently spaced apart, ensuring that no category is subsumed by another. The response options independently have a distinct probability of being selected by respondents across intervals of the latent trait (θ), which confirms the effectiveness of the scoring scale.

The item factor loadings ranged from +0.43 to +0.94, with Item 27 exhibiting the lowest and Item 16 the highest factor loading overall. Specifically, for the Participative Decision-Making Environment component, Item 9 had the lowest loading (+0.62) and Item 2 the highest (+0.84). For the Autonomy Supportive Environment component, Item 30 was the lowest (+0.41) and Item 25 was the highest (+0.83). For the Responsive Environment component, Item 10 was the lowest (+0.86) and Item 13 was the highest (+0.92). For the Professional Development Supportive Environment component, Item 17 was the lowest (+0.87) and Item 16 was the highest (+0.94). Finally, for the Facilitating Environment component, Item 21 was the lowest (+0.61) and Item 18 was the highest (+0.72).

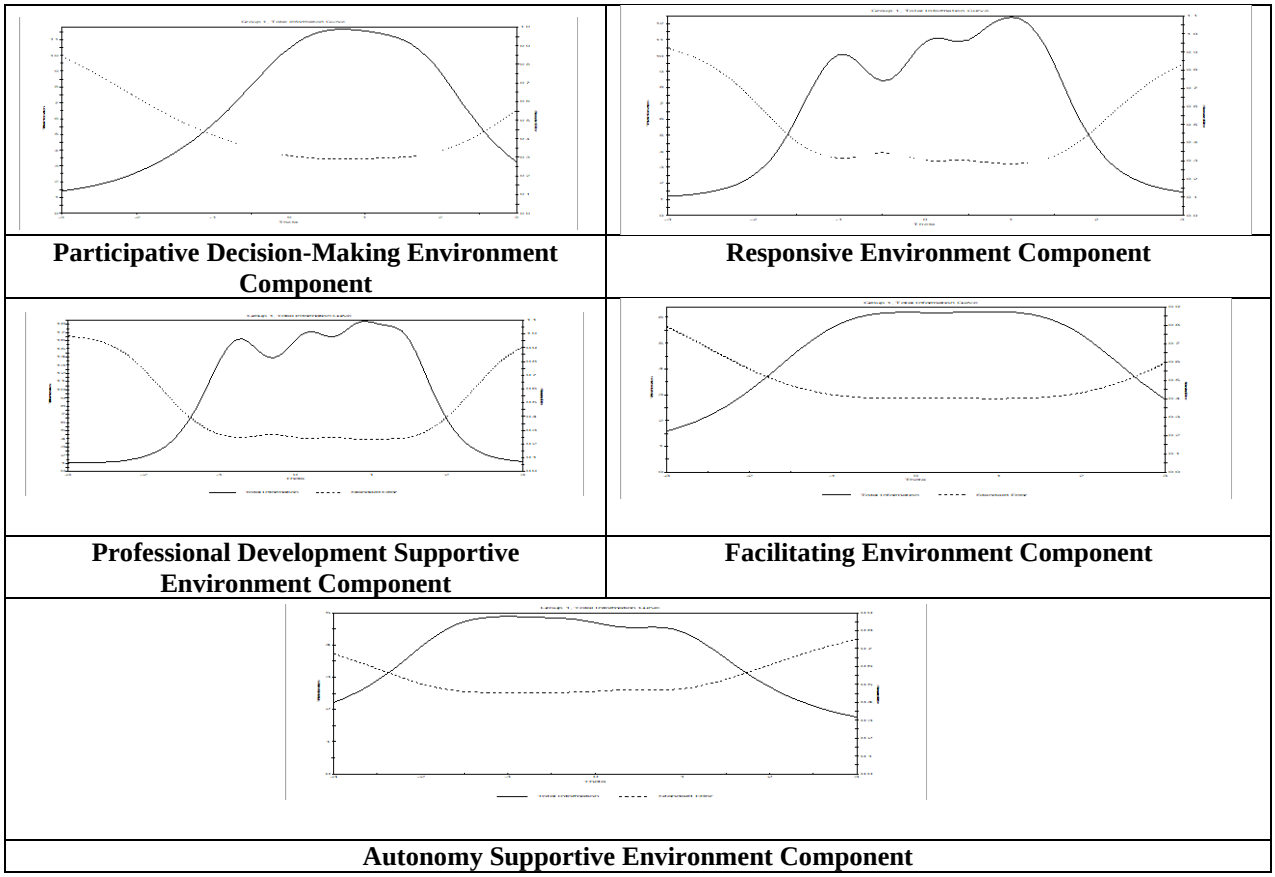

Figure 2. Test Information Function for the Components of the Teacher Empowering Structure Scale

To estimate the reliability of the scale, empirical or marginal reliability was employed. Considering that in Item Response Theory (IRT), the standard error of measurement (SEM) varies across respondents, the results indicated that the IRT-based marginal reliability coefficients were calculated as follows: 0.87 for the Participative Decision-Making Environment, 0.88 for the Responsive Environment, 0.90 for the Professional Development Supportive Environment, 0.82 for the Facilitating Environment, and 0.77 for the Autonomy Supportive Environment.

Discussion

The present study aimed to evaluate the psychometric properties of the Teacher Empowering Structure Scale (Altinkurt & Yilmaz, 2017) based on Classical Test Theory (CTT) and Item Response Theory (IRT) in an Iranian population. To determine the face validity of the scale, the opinions of 8 psychology experts were sought regarding item phrasing, comprehensibility,

desirability, and the avoidance of technical jargon. Revisions were made based on their feedback. Subsequently, to assess the Content Validity Ratio (CVR), the experts rated each item on a three-point continuum (not necessary = 0, useful = 1, essential = 2). The calculated CVR for the scale was 0.972. Furthermore, the Content Validity Index (CVI) yielded values of 0.96 for Clarity, 0.98 for Simplicity, and 0.96 for Relevance. Given that the minimum acceptable threshold for CVR and CVI, based on Lawshe's (1975) table for eight evaluators, is 0.75, it was concluded that the face and content validity of the Teacher Empowering Structure Scale were optimal.

Given the polytomous response pattern of the items, the Graded Response Model (GRM) was employed to examine the psychometric properties. The GRM is capable of analysing categorized responses along a continuum regarding attitudes or tendencies toward a phenomenon. IRTPRO software was used to calculate item parameters, verify the assumptions of local independence and one-dimensionality, and estimate the item information functions for each component of the scale. The assumption of local independence states that if the subject's ability level is considered the sole primary factor, the subjects' responses to the test items are statistically independent of one another. If this assumption holds, an examinee's performance on one item should not be influenced by their responses to other items (Sijtsma et al., 2002). Hambleton et al. (2011) argue that if the one-dimensionality assumption is met, local independence is also satisfied. In fact, by extracting a dominant factor, the covariance between variables approaches zero, and the relationship among items vanishes once the dominant factor is accounted for. According to Cai et al. (2016), items with chi-square values greater than 10 indicate potential local dependence. Similarly, Thissen et al. (2011) stated that marginal chi-square values exceeding 10 suggest a violation of local independence. A lack of local independence implies that the items do not solely load on a single dimension; rather, items with chi-square values greater than 10 form an additional dimension. Thus, local dependence indicates that a single dimension cannot adequately model the structure of those items.

Based on this assumption, the question arose: Are the responses to each item of the Teacher Empowering Structure Scale independent, such that answering or not answering one item does not affect the others? To test this, marginal chi-square values were calculated using Samejima's GRM via IRTPRO. Consistent with the criteria set by Cai et al. (2016) and Thissen et al. (2011), the

calculated marginal chi-square values for all items were less than 10, confirming that the assumption of local independence was satisfied.

Another core assumption of IRT is one-dimensionality, meaning there is a single dominant underlying factor that accounts for the item factor loadings and drives the selection of response categories. In other words, all test items should measure only one domain of ability or knowledge. Violating this assumption introduces bias in parameter estimation (Reckase, 2009). If the one-dimensionality assumption is not supported, the interpretation of test scores becomes ambiguous, comparisons between individuals' scores are flawed, and item parameter estimates are severely biased (De Champlain, 2010). Therefore, prior to applying the GRM, the one-dimensionality assumption of the scale was assessed using the NOHARM program. Multidimensional Item Response Theory (MIRT) analysis was utilized for this purpose. Three models (unidimensional, two-dimensional, and three-dimensional) were analysed and compared to infer the test's dimensionality. The results revealed that the log-likelihood for the unidimensional model was smaller than the other two models, indicating a better fit. Additionally, this model reached convergence with fewer iterations, further supporting its superiority.

When comparing the models, the chi-square statistics for the comparisons between the unidimensional and two-dimensional models, as well as the unidimensional and three-dimensional models, indicated significant differences. Furthermore, the Akaike Information Criterion (AIC) and the corrected Akaike Information Criterion (AICc) in the unidimensional model were larger compared to the two- and three-dimensional models. According to Tate's (2003) procedure, if the Root Mean Square Residual (RMSR) for the unidimensional model decreases by less than 10% compared to the two-dimensional model, the unidimensional model is deemed more appropriate. The results showed that the RMSR in the unidimensional model decreased by less than 10% compared to the two-dimensional model. Moreover, the calculated Tanaka index was 0.9301474 for the unidimensional model and 0.9215304 for the two-dimensional model. Based on these synthesized findings, the Teacher Empowering Structure Scale is unidimensional.

To further investigate one-dimensionality, Exploratory Factor Analysis (EFA) using Principal Axis Factoring was also conducted. In this method, the first factor must explain at least 20% of the variance, and the eigenvalue of the first factor is compared with the second (Reckase, 2009). Evaluating data suitability requires the Kaiser-Meyer-Olkin (KMO) index and Bartlett's Test of

Sphericity. The KMO index determines sample size adequacy; the closer to 1, the more adequate the data. Bartlett's test assesses whether the correlation matrix is significantly different from zero, tested via the chi-square distribution. If the null hypothesis (H_0) is rejected, factor analysis is justified. The analysis revealed an adequate KMO index and a significant Bartlett's test, confirming sample sufficiency and the factorability of the correlation matrix. The EFA with principal components showed that the first extracted factor explained 22.13% of the variance, followed by the second (11.20%), third (11.09%), fourth (8.40%), and fifth (7.06%), cumulatively explaining 59.88% of the variance. According to Reckase (2009) and Kirisci et al. (2001), the presence of one dominant factor alongside several minor factors supports the application of a unidimensional model.

An appropriate scale's items must be able to differentiate between individuals with high and low levels of the latent trait. To address this, item discrimination parameters must be examined. In the GRM, the discrimination parameter indicates the extent to which an item relates to the latent trait level, and how much the probability of selecting a specific category increases with a one-unit increase in the latent trait. In other words, it reflects how rapidly the expected item scores change with variations in the latent trait (θ). According to Baker and Kim (2004), discrimination parameters below 0.65 are considered low, 0.65 to 1.34 as moderate, and 1.35 or higher as high. Baker (2001) suggests that a good item in a normal ogive model should have a discrimination parameter (a) of at least 0.65, and items with low discrimination ($a < 0.65$) should be excluded from the analysis.

The evaluation of item fit indices based on the GRM showed that the calculated chi-square significance level for all items was greater than 0.05, indicating good fit. The item discrimination parameters for the overall scale ranged from +0.77 to +5.29. Specifically, these ranged from +1.34 to +2.62 for the Participative Decision-Making Environment; +0.77 to +2.25 for the Autonomy Supportive Environment; +2.83 to +5.29 for the Responsive Environment; +2.99 to +4.44 for the Professional Development Supportive Environment; and +1.31 to +1.74 for the Facilitating Environment. Furthermore, the first threshold parameter ranged from -2.66 to +0.39, the second from -1.99 to +0.78, the third from -1.02 to +2.24, and the fourth from +0.33 to +2.44. Consequently, the Category Response Curves (CRCs) were well-ordered, with clearly distinguishable distributions. The curves lacked flatness and unnecessary intertwining. The

thresholds for all response categories were sufficiently spaced apart, meaning no category was subsumed by another. The response options independently possessed a distinct probability of selection across latent trait (θ) intervals, thereby confirming the effectiveness of the scoring scale.

The item factor loadings ranged from +0.43 to +0.94, with Item 27 exhibiting the lowest and Item 16 the highest factor loading. Specifically, for the Participative Decision-Making Environment component, Item 9 had the lowest loading (+0.62) and Item 2 the highest (+0.84). For the Autonomy Supportive Environment component, Item 30 was the lowest (+0.41) and Item 25 the highest (+0.83). For the Responsive Environment component, Item 10 was the lowest (+0.86) and Item 13 the highest (+0.92). For the Professional Development Supportive Environment component, Item 17 was the lowest (+0.87) and Item 16 the highest (+0.94). Finally, for the Facilitating Environment component, Item 21 was the lowest (+0.61) and Item 18 the highest (+0.72). The item information functions yielded the highest amount of information around the average latent trait (θ), indicating that the items possess an optimal level of informativeness.

To estimate the reliability, empirical or marginal reliability was utilized. Given that the standard error of measurement varies across examinees in the Item Response Theory (IRT) approach, the results indicated that the IRT-based marginal reliability coefficients were 0.87 for the Participative Decision-Making Environment, 0.88 for the Responsive Environment, 0.90 for the Professional Development Supportive Environment, 0.82 for the Facilitating Environment, and 0.77 for the Autonomy Supportive Environment.

Like other scientific studies, the present research encountered several limitations that should be taken into consideration in future investigations. One such limitation was conducting the study exclusively among teachers in the city of Bandar Abbas, which may have influenced the findings. The failure to account for variables such as the socioeconomic and cultural status of the teachers, as well as the self-report nature of the questionnaire, constituted additional limitations that could restrict the generalizability of the findings to the broader population.

Therefore, it is recommended that future studies replicate the Teacher Empowering Structure Scale across other geographical regions of Iran (aside from Bandar Abbas) to ensure greater confidence in generalizing the results. Moreover, considering the socioeconomic and cultural status of participants in future research can provide further empirical evidence regarding the applicability

of this scale within the Iranian population while facilitating safer generalizations. Finally, given the findings of the present study confirming the psychometric properties (adequate validity and reliability) of the Teacher Empowering Structure Scale, this instrument is deemed valid and reliable for use by domestic researchers. It can provide a suitable framework for assessing the empowering structure among the teacher population in Bandar Abbas and serve as a preliminary step for organizational planning to enhance this structure.

Data availability statement

The original contributions presented in the study are included in the article; further inquiries can be directed to the corresponding author.

Ethics statement

The studies involving human participants were reviewed and approved by the Ethics Committee of Islamic Azad University. All participants provided their written informed consent to participate in this study.

Author contributions

Masoume Bahamin served as the principal investigator and drafted the manuscript; Nooshin Taghinezhad acted as the primary supervisor and contributed to the study design and data analysis; Abdulhadi Samavi and Kobra Hajializadeh contributed as advising supervisors and reviewed the manuscript. All authors approved the submitted version.

Funding

This research was conducted as part of a Ph.D. dissertation and received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

Addi-Raccah, A. (2009). Between teachers' empowerment and supervision: A comparison of school leaders in the 1990s and the 2000s. *Management in Education, 23*(4), 161-167.

Akgun, N. (2021). The relation between learning-centered leadership and structural empowerment of teachers. *International Online Journal of Educational Sciences, 13*(4), 965-983.

Altinkurt, Y., & Iliman Puskulluoglu, E. (2017). Development of Teachers' Structural Empowerment Scale (TSES): A validity and reliability study. *Üniversitepark Bülten, 6*(1), 118-132. https://doi.org/10.22521/unibulletin.2017.61.10

Altinkurt, Y., & Yilmaz, K. (2013). Development of organizational power scale at schools: A reliability and validity study. *E-International Journal of Educational Research, 4*(4), 1-17.

Altınkurt, Y., Türkkaş Anasız, B., & Ergin Ekinci, C. (2016). Öğretmenlerin yapısal ve psikolojik güçlendirilmeleri ile örgütsel vatandaşlık davranışları arasındaki ilişki. *Eğitim ve Bilim, 41*(187), 79-96.

Al-Yaseen, W. S., & Al-Musaileem, M. Y. (2015). Teacher empowerment as an important component of job satisfaction: A comparative study of teachers' perspectives in Al-Farwaniya District, Kuwait. *Compare: A Journal of Comparative and International Education, 45*(6), 863-885.

Baker, F. B. (2001). *The basics of item response theory*. ERIC Clearinghouse on Assessment and Evaluation.

Baker, F. B., & Kim, S. H. (2004). *Item response theory: Parameter estimation techniques*. CRC Press.

Balkar, B. (2015). Defining an empowering school culture (ESC): Teacher perception. *Issues in Educational Research, 25*(3), 205-226.

Bogler, R., & Nir, A. E. (2012). The importance of teachers' perceived organizational support to job satisfaction: What empowerment got to do with it. *Journal of Educational Administration, 50*(3), 287-306.

Bogler, R., & Somech, A. (2004). Influence of teacher empowerment on teachers' organizational commitment, professional commitment and organizational citizenship behavior in schools. *Teaching and Teacher Education, 20*(3), 277-289.

Cai, L., Choi, K., Hansen, M., & Harrell, L. (2016). Item response theory. *Annual Review of Statistics and Its Application, 3*(1), 297-321.

Comrey, A. L., & Lee, H. B. (2013). *A first course in factor analysis*. Psychology Press.

De Champlain, A. F. (2010). A primer on classical test theory and item response theory for assessments in medical education. *Medical Education, 44*(1), 109-117.

Dee, J. R., Henkin, A. B., & Duemer, L. (2003). Structural antecedents and psychological correlates of teacher empowerment. *Journal of Educational Administration, 41*(3), 257–277.

DeMars, C. (2010). *Item response theory*. Oxford University Press.

Fandiño, Y. J. (2010). Research as a means of empowering teachers in the 21st century. *Educación y Educadores, 13*(1), 109-124.

Hambleton, R. K., van der Linden, W. J., & Wells, C. S. (2011). IRT models for the analysis of polytomously scored data: Brief and selected history of model building advances. In M. Nering & R. Ostini (Eds.), *Handbook of polytomous item response theory models* (pp. 21-42). Routledge.

Hooman, H. A. (2011). *Multivariate data analysis in behavioral research* (3rd ed.). Peyk-e Farhang Publications. [In Persian].

In'am, A. (2015). Mathematics teachers' response towards teacher empowerment at junior high school in Indonesia. *American Journal of Educational Research, 3*(4), 439-445.

Jiang, S., Wang, C., & Weiss, D. J. (2016). Sample size requirements for estimation of item parameters in the multidimensional graded response model. *Frontiers in Psychology, 7*, Article 109.

Jiang, Y., Li, P., Wang, J., & Li, H. (2019). Relationships between kindergarten teachers' empowerment, job satisfaction, and organizational climate: A Chinese model. *Journal of Research in Childhood Education, 33*(2), 257-270.

Kanter, R. M. (1993). *Women and men of the corporation* (2nd ed.). Basic Books.

Kiani, S., Kadivar, P., Keramati, H., & Geramipour, M. (2022). Psychometric properties of the human agency scale: Classical test theory and graded response model. *Educational Measurement, 13*(47), 113-144. https://doi.org/10.22054/jem.2021.58309.2134 [In Persian].

Kimwarey, M. C., Chirure, H. N., & Omondi, M. (2014). Teacher empowerment in education practice: Strategies, constraints and suggestions. *IOSR Journal of Research & Method in Education*, *4*(2), 51-56.

Kirisci, L., Hsu, T. C., & Yu, L. (2001). Robustness of item parameter estimation programs to assumptions of unidimensionality and normality. *Applied Psychological Measurement, 25*(2), 146-162.

Kline, R. B. (2023). *Principles and practice of structural equation modeling* (5th ed.). Guilford Publications.

Lawshe, C. H. (1975). A quantitative approach to content validity. *Personnel Psychology, 28*(4), 563–575.

Le, D. (2013). *Applying item response theory modeling in educational research* [Master's thesis, Iowa State University]. Iowa State University Digital Repository.

Morales-Thomas, M. (2015). *Practices that promote parent engagement in an urban elementary school: A phenomenological study of Latino parents of English language learners* [Doctoral dissertation, Boston College]. ProQuest Dissertations and Theses Global.

Nering, M., & Ostini, R. (Eds.). (2011). *Handbook of polytomous item response theory models*. Taylor & Francis.

Ökmen, A. (2018). *Öğretmen güçlendirmeye ilişkin lisede görev yapan öğretmenlerin algıları* [The perceptions of high school teachers regarding teacher empowerment] [Unpublished master's thesis]. Yıldız Teknik University.

Özer, N., & Dönmez, B. (2013). Reassessing the psychometric properties of the Turkish version of Enabling School Structure Scale. *Pegem Eğitim ve Öğretim Dergisi, 3*(4), 57–68. https://doi.org/10.14527/C3S4M4

Ozkan Hidiroglu, Y., & Tanriogen, A. (2020). Development of Teachers' Empowerment Scale: A validity and reliability study. *International Journal of Assessment Tools in Education, 7*(4), 753–772. https://doi.org/10.21449/ijate.693398

Pelit, E., Ozturk, Y., & Arslanturk, Y. (2011). The effects of employee empowerment on employee job satisfaction: A study on hotels in Turkey. *International Journal of Contemporary Hospitality Management, 23*(6), 784-802.

Reckase, M. D. (2009). *Multidimensional item response theory*. Springer.

Rice, E. M., & Schneider, G. T. (1994). A decade of teacher empowerment: An empirical analysis of teacher involvement in decision making, 1980-1991. *Journal of Educational Administration, 32*(1), 43–58.

Rinehart, J. S., & Short, P. M. (1994). Job satisfaction and empowerment among teacher leaders, reading recovery teachers, and regular classroom teachers. *Education, 114*(4), 570–580.

Rinehart, J. S., Short, P. M., Short, R. J., & Eckley, M. (1998). Teacher empowerment and principal leadership: Understanding the influence process. *Educational Administration Quarterly, 34*(4), 630-649.

Sharma, S. A. (2014). Coaching learning: Perspectives on teaching and learning. *Michigan Reading Journal, 46*(2), 57-74.

Sijtsma, K., & Molenaar, I. W. (2002). *Introduction to nonparametric item response theory* (Vol. 5). SAGE Publications.

Somech, A. (2005). Teachers' personal and team empowerment and their relations to organizational outcomes: Contradictory or compatible constructs? *Educational Administration Quarterly, 41*(2), 237–266.

Spreitzer, G. M. (1995). Psychological empowerment in the workplace: Dimensions, measurement and validation. *Academy of Management Journal, 38*(5), 1442-1465.

Tate, R. (2003). A comparison of selected empirical methods for assessing the structure of responses to test items. *Applied Psychological Measurement, 27*(3), 159–203.

Thissen, D., Cai, L., & Bock, R. D. (2011). The nominal categories item response model. In M. Nering & R. Ostini (Eds.), *Handbook of polytomous item response theory models* (pp. 43-75). Routledge.

Wilson, S., & Coolican, M. J. (1996). How high and low self-empowered teachers work with colleagues and school principals. *Journal of Educational Thought, 30*(2), 99-118.

Yin, H. B., Jin, Y. L., & Lee, C. K. (2009). The impact of trust and empowerment culture on curriculum reform. *Journal of Capital Normal University, 1*, 125–132.

Zimmerman, M. A. (2000). Empowerment theory: Psychological, organizational and community levels of analysis. In J. Rappaport & E. Seidman (Eds.), *Handbook of community psychology* (pp. 43-63). Kluwer Academic/Plenum.