Critical Public Health

ISSN: 0958-1596 (Print) 1469-3682 (Online) Journal homepage: www.tandfonline.com/journals/ccph20

Psychometric validation of physical education teachers' self-efficacy scale

Veli Başal, Hüseyin Ünlü, Metin Yüceant & Monira I. Aldhahi

To cite this article: Veli Başal, Hüseyin Ünlü, Metin Yüceant & Monira I. Aldhahi (2026) Psychometric validation of physical education teachers' self-efficacy scale, Critical Public Health, 36:1, 2656069, DOI: 10.1080/09581596.2026.2656069

To link to this article: https://doi.org/10.1080/09581596.2026.2656069

CRITICAL PUBLIC HEALTH
2026, VOL. 36, NO. 1, 2656069
https://doi.org/10.1080/09581596.2026.2656069

RESEARCH ARTICLE

OPEN ACCESS Check for updates

Psychometric validation of physical education teachers' self-efficacy scale

Veli Başal[a], Hüseyin Ünlü[a], Metin Yüceant[a] and Monira I. Aldhahi[b]

[a]Faculty of Sports Sciences, Aksaray University, Aksaray, Turkey; [b]Department of Rehabilitation Sciences, College of Health and Rehabilitation Sciences, Princess Nourah bint Abdulrahman University, Riyadh, Saudi Arabia

ABSTRACT

The aim of this study is to develop a valid and reliable measurement instrument, grounded in Bandura's social cognitive theory, to comprehensively assess the self-efficacy of physical education teachers. This psychometric and cross-sectional methodological study was conducted using a survey-based scale development design. The initial item pool, generated through a comprehensive literature review, was evaluated based on expert opinions, and content validity was established following a pilot study. The study sample consisted of 419 physical education teachers from various regions of Turkey. Construct validity was examined using item-total correlations, exploratory factor analysis (EFA), and confirmatory factor analysis (CFA). Reliability was assessed using Cronbach's alpha coefficients. Item-total correlation coefficients were found to be above 0.45. EFA results revealed a four-factor structure explaining 78.799% of the total variance: teaching process management, student motivation, field expertise, and professional development. CFA fit indices indicated a good model fit. The overall Cronbach's alpha coefficient of the scale was 0.824, while subscale coefficients ranged between 0.742 and 0.822. Compared to existing instruments, the developed scale may contribute to the literature by providing a comprehensive tool that reflects the socio-cultural and institutional realities of education, focuses on in-service teachers, and holistically captures the unique dimensions of physical education teaching. The Physical Education Teachers' Self-Efficacy Scale can be considered a valid and reliable instrument for assessing professional competence, informing in-service training programs, and supporting future research.

ARTICLE HISTORY
Received 1 March 2026
Accepted 2 April 2026

KEYWORDS
Self-efficacy; physical education teacher; scale development; validity; reliability

Introduction

Education is a fundamental process that enables individuals to actively participate in social life by developing cognitive, affective, and psychomotor skills. Teachers, who are at the center of this process, play a key role in achieving the goals of the education system (Guskey & Passaro, 1994). Teachers' professional competencies, attitudes, and beliefs directly influence the quality of education they provide and are among the key determinants of student achievement (Yenen, 2022). In this context, one of the most important psychological constructs shaping teacher behavior is self-efficacy (Karaiskos et al., 2024).

According to Bandura's social cognitive theory, self-efficacy is an individual's belief in their own capability to successfully organize and execute a specific task (Bandura, 1977). This belief directly affects the amount of effort individuals expend, their persistence, and their ability to cope with stress. High self-efficacy is associated with greater motivation and better performance in challenging tasks (Bandura, 1997). In the educational context, teacher self-efficacy has critical effects on teacher satisfaction, burnout, student engagement, and achievement (Tschannen-Moran & Hoy, 2001). In physical education, self-efficacy is a key factor that enhances teachers' effectiveness in directing physical activity, increasing student

CONTACT Monira I. Aldhahi mialdhahi@pnu.edu.sa
Supplemental data for this article can be accessed online at https://doi.org/10.1080/09581596.2026.2656069.

motivation, and promoting healthy lifestyle habits (Zee & Koomen, 2016). The unique requirements of this field lead to the inadequacy of general teacher self-efficacy scales.

The measurement of teacher self-efficacy has long been a topic in academic literature. Although Bandura's General Self-Efficacy Scale has been used across different domains, the multidimensional nature of the teaching profession has necessitated the development of profession-specific instruments (Bandura, 2006). In this context, the Teacher Self-Efficacy Scale, which includes dimensions such as classroom management, instructional strategies, and student engagement, is widely used (Tschannen-Moran & Hoy, 2001). However, instability in its factor structure and its inability to reflect contextual differences specific to fields such as physical education are considered important limitations in literature.

Having a high level of self-efficacy among physical education teachers is critically important for student development. Indeed, a positive relationship has been found between teacher self-efficacy and students' exercise motivation and performance (Karaiskos et al., 2024). In addition, teachers who demonstrate high self-efficacy in areas such as motivation, equipment use, and adapting instruction to diverse needs have been shown to increase the effectiveness of lessons (Humphries et al., 2012).

The number of scales developed for this field is limited. The PETES scale developed by Humphries et al. (2012) addresses dimensions such as instruction, student motivation, and subject matter knowledge; however, it has limitations in terms of inclusive practices and cultural adaptability. Pan (2014) proposed a model focusing on student motivation and classroom climate, but did not include professional development dimensions. The scale developed by Woodcock et al. (2025), although based on Bandura's Triadic Reciprocal Determinism model, addresses the practical challenges of developing countries only to a limited extent.

Studies conducted in Türkiye have also examined physical education teachers' self-efficacy from different perspectives. Ünlü et al. (2008) developed a scale assessing professional competencies, Esentürk et al. (2019) introduced a scale measuring interpersonal self-efficacy, and Sayan et al. (2022) evaluated pre-service teachers' self-efficacy related to special teaching methods. In addition, it has been found that physical education teachers demonstrate innovative attitudes but need to further develop their self-efficacy (Yalvuç & Karli, 2022). These findings indicate that self-efficacy is directly related to teacher education policies and professional development practices.

When the existing literature is considered as a whole, there is a clear need for a comprehensive and culturally adapted measurement tool specific to physical education teachers. Most existing scales are designed for general teachers or pre-service teachers and do not adequately reflect the real working conditions of in-service physical education teachers. Contextual factors such as facility limitations, large class sizes, differences in equipment, and cultural expectations have been addressed only to a limited extent. Moreover, some scales being culture-specific limits their adaptability, while others fail to address self-efficacy in a holistic manner, thereby restricting content validity (Block et al., 2013; Braksiek, 2022; Eroğlu & Ünlü, 2015). It is also evident that Bandura's principle of reciprocal determinism, as well as contemporary dimensions such as inclusive education, technology integration, and professional development, have not been sufficiently incorporated.

This study aims to develop a comprehensive self-efficacy scale for physical education teachers in order to address this gap. The scale consists of four main dimensions: teaching process management, student motivation, field expertise, and professional development. These dimensions are directly related to Bandura's theoretical framework and have not been sufficiently integrated in existing instruments. It is expected that the developed measurement tool will contribute to supporting teachers' professional development, improving teacher education programs, and providing evidence-based input for educational policies. Furthermore, it has the potential to be adapted and used across different disciplines. In this context, the Physical Education Teacher Self-Efficacy Scale is expected to make a meaningful contribution to the literature due to its original structure and multidimensional nature.

Material and methods

Study design

This psychometric and cross-sectional methodological study focused on scale development and was conducted using a survey-based design. The research process consisted of item pool generation, expert

review, pilot testing, and validity–reliability analyses. Ethical approval was obtained from the Human Research Ethics Committee of Aksaray University (decision dated 18.11.2025, numbered 2025-453), and all procedures were carried out in accordance with the Declaration of Helsinki. Written informed consent was obtained from all participants; participation was voluntary, and data were processed in line with confidentiality principles.

The study is a survey research aiming to determine physical education teachers' self-efficacy perceptions; survey designs are approaches that describe the current situation without intervention (Karasar, 2009).

Study participants

The population of this scale development study consisted of physical education teachers working in public and private institutions affiliated with the Ministry of National Education in Türkiye. The sample size was calculated using G*Power 3.1.9.4 based on assumptions of 95% power (1-β = 0.95), α = 0.05, and a medium effect size (f = 0.25), and a minimum of 250 participants was found to be sufficient. In the study, 419 participants were divided into two groups: a pilot sample (n = 103), which was used for item pool refinement, and a final sample (n = 316), which was used for exploratory and confirmatory factor analyses. Conducting analyses on separate datasets reduced the risk of bias and provided cross-validation.

A maximum variation sampling approach was adopted in participant selection, and teachers with diverse characteristics in terms of gender, professional experience, school type, geographical location, and educational level were included. In the literature, a sample size of 300 or more is considered a 'good' level, and it is recommended that the sample be at least five times the number of items. Accordingly, the sample size of the study is considered adequate and acceptable (Tabachnick & Fidell, 2013; Tavşancıl, 2014).

According to Table 1, the mean age of the participants was 35.99 years. Regarding gender distribution, 139 participants (33.2%) were female, and 280 (66.8%) were male. The mean professional experience was 11.78 years. In terms of educational background, 285 participants (68.0%) held a bachelor's degree, 132 (31.5%) a master's degree, and 2 (0.5%) a doctoral degree. With respect to the type of school in which they worked, 292 participants (69.7%) were employed in middle schools, while 127 (30.3%) taught in high schools. Considering the location of the schools, 126 participants (30.1%) worked in villages, 88 (21.0%) in districts, and 205 (48.9%) in provincial centers.

Development of the draft scale and ensuring content validity

At the initial stage of the scale development process, relevant self-efficacy scales were reviewed based on a comprehensive literature search, and a conceptual framework was established, resulting in a 56-item item pool. To evaluate the items in terms of content and language, the opinions of seven experts with extensive experience in their respective fields were obtained (2 in educational sciences, 1 in measurement and evaluation, 1 in language, and 3 in sports sciences). The experts were asked to rate the items using a 5-point Likert-type scale and to provide qualitative feedback and suggestions. All experts had at

Table 1. Demographic characteristics of participants.

Personal information	Group	N	%
Age	$\bar{X}$ = 35.99		
Gender	Female	139	33.2
	Male	280	66.8
Years of service	$\bar{X}$ = 11.78		
	Bachelor's	285	68.0
Education level	Master's degree	132	31.5
	Doctorate	2	0.5
School level you work at	Middle school	292	69.7
	High School	127	30.3
	Village	126	30.1
Location of the school you work at	District	88	21.0
	City center	205	48.9

Denotes data are presented as Frequency (N), and Percentage (%).

least one published scale development study and a minimum of 10 years of academic experience in their fields.

Expert evaluations were analyzed using content validity indices (CVR and I-CVI). According to Lawshe (1975), the minimum CVR value for seven experts is 0.99, and 20 items falling below this threshold were removed. The remaining items had I-CVI values ranging between 0.86 and 1.00, and the S-CVI/Ave value of the scale was found to be 0.94, indicating that content validity was achieved (Lynn, 1986; Polit & Beck, 2006). As a result, 20 items were removed, 5 new items were added, and a 41-item draft scale was developed.

Data collection tools

Two instruments were used for data collection in this study. The first was a Personal Information Form consisting of seven questions designed to determine the socio-demographic characteristics of physical education teachers. The second was the Physical Education Teachers' Self-Efficacy Scale, developed to measure teachers' self-efficacy levels. This scale consists of 24 items and uses a 5-point Likert-type format. Data were collected through online forms between November and December 2025.

Data analysis

In the validity and reliability process of the scale, item-total correlations, the Kaiser–Meyer–Olkin (KMO) coefficient, and Bartlett's Test of Sphericity were used to evaluate data suitability. Exploratory factor analysis (EFA) was conducted, and confirmatory factor analysis (CFA) was applied to validate the factor structure. Internal consistency was assessed using Cronbach's alpha coefficient, while relationships between variables were examined using Pearson correlation analysis. In the EFA, Principal Component Analysis and Varimax rotation methods were employed, and all analyses were performed using SPSS 21 and AMOS 21 software. Prior to factor analysis, assumptions of normality, linearity, multicollinearity, and outliers were tested. It was determined that VIF values were below 5 and that no outliers were present in the dataset.

Results

This section presents the results obtained from the validity and reliability analyses of the scale used in this study. Validity refers to the extent to which the measurement tool accurately measures the targeted characteristic, whereas reliability refers to the repeatability of the same measurements on the same individuals under similar conditions (Büyüköztürk et al., 2022).

Validity analysis

Based on expert opinions, a pilot study was initiated with 103 physical education teachers using a draft scale of 41 items. While the literature indicates that 30–40 participants are sufficient for pilot studies, it is emphasized that the total sample size should not fall below 100 (Beaton et al., 2000; Foster, 2001). During the pilot application process, participants were asked for their opinions on the items. Based on the feedback obtained and the statistical analyses performed, 12 items were removed from the pool. Thus, a final item pool consisting of four sub-dimensions representing the self-efficacy of physical education teachers and 29 items related to these dimensions was created.

Construct validity

To determine the construct validity of the 41 items in the pilot form of the scale, total test correlation coefficients were calculated for each item.

Item-total correlations were examined in the draft form of the Physical Education Teachers' Self-Efficacy Scale, and the cutoff point was set at .45 (Büyüköztürk, 2023). As shown in Table 2, Items 6, 9, 10, 13, 15, 18, 22, 23, 28, 29, 31, and 36 were below this value and were therefore removed from the scale. After this adjustment, the number of items in the scale was 29.

It is stated that for an item to remain on the scale during the CFA process, there must be a difference of at least .10 between the load values in different factors (Büyüköztürk, 2023). In line with this criterion, five items (7, 16, 25, 34, and 39) that showed a high level of interdependence with other factors were removed from the scale. Thus, the scale was finalized with 24 items. There are no reverse items in the scale; the minimum score that can be obtained from the scale is 24, and the maximum score is 120.

Exploratory factor analysis (EFA)

The KMO coefficient and Bartlett's Sphericity Test were applied to evaluate the suitability of the draft scale for factor analysis. During the scale development process, it is stated that the KMO value must be at least 0.60 and the chi-square value obtained from Bartlett's Sphericity Test must be significant to perform factor analysis (Büyüköztürk, 2023). The KMO and Bartlett's Sphericity Test findings calculated in this study are shown in Table 3.

As shown in Table 3, the Kaiser–Meyer–Olkin (KMO) measure of sampling adequacy was 0.863, and Bartlett's test yielded a χ^2 value of 2862.243, with a significance level of .000 ($p < 0.001$). In the literature, KMO values of 0.50 are considered 'acceptable,' approximately 0.60 'mediocre,' 0.70 'good,' 0.80 'very good,' and values above 0.90 indicate 'excellent' sampling adequacy (Şeker et al., 2004). These results indicate that the sample used in this study was suitable for factor analysis. Principal Component Analysis was employed as the factor extraction method, and the number of factors was determined using the eigenvalue >1 criterion. Rotation was performed using the Varimax method, which ensures that factors are orthogonal and independent, thereby enhancing interpretability (Büyüköztürk, 2023; Tabachnick & Fidell, 2013).

The factor loadings and variance explanation ratios of the scale draft are presented in Table 4.

According to the exploratory factor analysis (EFA) results presented in Table 4, four factors with eigenvalues above 1 emerged in the draft scale. The total variance explained by these factors was calculated as 78.799%. Based on the CFA findings, the factor structures in the draft scale and the items belonging to these factors are shown in Table 5.

According to Table 5, the scale consists of four factors, each composed of six items. The factor loadings for the first factor range from 0.496 to 0.673, for the second factor from 0.488 to 0.723, for the third factor from 0.574 to 0.766, and for the fourth factor from 0.638 to 0.733. The scale is named the 'Physical Education Teachers' Self-Efficacy Scale' because it was developed to measure the professional self-efficacy beliefs of physical education teachers.

Table 2. Item-total test correlation coefficients of the scale draft.

Item No	r	Item No	r	Item No	r
1	0.673	15	0.275	29	0.286
2	0.581	16	0.497	30	0.763
3	0.496	17	0.633	31	0.417
4	0.532	18	0.354	32	0.766
5	0.602	19	0.689	33	0.698
6	0.392	20	0.652	34	0.475
7	0.482	21	0.588	35	0.638
8	0.547	22	0.296	36	0.412
9	0.405	23	0.327	37	0.701
10	0.442	24	0.574	38	0.733
11	0.488	25	0.513	39	0.515
12	0.525	26	0.569	40	0.674
13	0.409	27	0.594	41	0.689
14	0.723	28	0.379		

Table 3. KMO and Bartlett's Sphericity Test results.

KMO Coefficient		0.863
Bartlett's Sphericity Test	Chi-square	2862.243
	df	222
	p	.000

Table 4. Factor eigenvalues and variance explanation ratios.

Factors	Factor eigenvalues	Variance %	Total variance %
1	9.135	40.614	40.614
2	5.241	23.301	63.915
3	2.014	8.954	72.869
4	1.334	5.930	78.799

Table 5. Exploratory factor analysis (EFA) results.

			Factors			
Factor name	Old item number	New item number	1	2	3	4
Teaching Process Management	1	1	0.673			
	5	2	0.602			
	2	3	0.581			
	8	4	0.547			
	4	5	0.532			
	3	6	0.496			
Student Motivation	14	7		0.723		
	19	8		0.689		
	20	9		0.652		
	17	10		0.633		
	12	11		0.525		
	11	12		0.488		
Field Expertise	32	13			0.766	
	30	14			0.763	
	27	15			0.594	
	21	16			0.588	
	26	17			0.569	
	24	18			0.574	
Professional Development	38	19				0.733
	37	20				0.701
	33	21				0.698
	41	22				0.689
	40	23				0.674
	35	24				0.638

Physical education has unique learning environments, objectives, and implementation conditions that require specific competency areas not fully captured by general teacher self-efficacy scales. While the Teacher Self-Efficacy Scale in the literature (Tschannen-Moran & Hoy, 2001) focuses on general dimensions such as classroom management, student engagement, and instructional strategies, physical education teaching involves distinctive competencies such as physical interaction, space management, equipment limitations, safety, inclusive practices, and technology integration (Humphries et al., 2012; Pan, 2014). Moreover, existing international scales were developed in Western cultural contexts and do not fully cover all aspects of physical education. National measurement tools, on the other hand, generally focus on limited sub-dimensions (Esentürk et al., 2019; Sayan et al., 2022).

To address this gap, the developed scale consists of items reflecting the daily practices of physical education teachers and the socio-cultural conditions of the Ministry of National Education system in Türkiye. The naming of the sub-dimensions is based on the relevant academic literature. Accordingly, the factors represent teaching process management, student motivation, field expertise, and professional development. Therefore, the scale is considered a psychometric instrument specific to the field of physical education, independent from general teacher self-efficacy scales.

Confirmatory factor analysis (CFA)

In this stage, confirmatory factor analysis (CFA) was conducted on the scale structure consisting of four latent variables (dimensions)—teaching process management, student motivation, field expertise, and professional development—and 24 observed variables (items) representing these dimensions. The CFA was performed on the final sample of 316 participants, which was completely independent of the pilot sample, to validate the factor structure obtained from the exploratory factor analysis (EFA).

To evaluate model fit, multiple fit indices (χ^2/df, GFI, AGFI, NFI, CFI, IFI, TLI, and RMSEA) were examined collectively. This approach mitigates the limitations of relying on a single fit index and provides a more comprehensive assessment of model fit (Hu & Bentler, 1999).

Table 6. Internal consistency coefficients and correlation values for the total scale and factors.

	α	TPM	SM	FE	PD
TPM	0.742	1	0.617*	0.704*	0.588*
SM	0.796		1	0.644*	0.756*
FE	0.814			1	0.818*
PD	0.822				1
Total	0.824				

*$p < 0.01$; TPM: Teaching Process Management; SM: Student Motivation; FE: Field Expertise; PD: Professional Development.

Among these indices, CFI, TLI, and RMSEA are considered key criteria for determining model adequacy. An RMSEA value ≤ 0.08 indicates good fit, while a value ≤ 0.05 represents excellent fit. Similarly, CFI and TLI values above 0.85 are indicative of a good fit.

In the initial stage, the fit indices of the measurement model were examined, but some values were observed to be close to threshold levels. To improve model fit, modification indices were considered. Upon reviewing these indices, high error covariances were identified between certain items that belonged to the same factor and exhibited conceptual similarity.

Only a limited number of error covariances were specified between item pairs within the same sub-dimension, ensuring that the modifications were theoretically meaningful and consistent with the nature of the measurement model. These adjustments were applied carefully so as not to alter the overall factor structure or affect the relationships among the latent variables. Following these modifications, noticeable improvements in model fit indices were observed.

When examining the fit index of the scale, x^2/df ($2.897 < 3$), AGFI ($0.866 > 0.85$), NFI ($0.891 > 0.85$), CFI ($0.901 > 0.85$), IFI ($0.905 > 0.85$), TLI ($0.912 > 0.85$), and RMSEA ($0.059 < 0.08$) indices show a good level of fit.

The ability of a scale to produce similar and consistent results across different measurements indicates the reliability of that scale (Ercan & Kan, 2004). Cronbach's Alpha Reliability Coefficient was used to evaluate the reliability of the scale developed in this study. This coefficient, which provides important information about internal consistency and is accepted as evidence of reliability, is one of the most frequently used methods in the fields of education and psychology. It is a basic reliability indicator known to the vast majority of researchers interested in the field of measurement and evaluation (Kartal & Dirlik, 2016).

The overall reliability values of the Physical Education Teachers' Self-Efficacy Scale and the internal consistency and correlation coefficients for the subscales are shown in Table 6.

Cronbach's Alpha (α) internal consistency coefficients for the entire scale and its subscales were found to be above 0.70. In addition, positive correlations were found between the subscales and between the subscales and the overall scale. A Cronbach's Alpha coefficient in the range of 0.80–1.00 indicates high reliability, 0.60–0.79 indicates acceptable reliability, 0.40–0.60 indicates low reliability, and 0.00–0.39 indicates insufficient reliability (Büyüköztürk, 2023). In this study, the overall Cronbach's alpha coefficient of the scale was 0.824, which indicates that the scale has high internal consistency and reliability.

Convergent and discriminant validity

To evaluate the construct validity of the scale more comprehensively, convergent and discriminant validity analyses were conducted. For this purpose, the Average Variance Extracted (AVE) and Composite Reliability (CR) values were calculated for each factor. Convergent validity is considered adequate when AVE values exceed 0.50 and CR values exceed 0.70 (Fornell & Larcker, 1981). The analyses indicated that all sub-dimensions had AVE values above 0.50 and CR values above 0.70, demonstrating adequate convergent validity. Discriminant validity was assessed by comparing the square root of each factor's AVE with the correlations between factors. The results showed that, for all factors, the square root of the AVE exceeded the inter-factor correlations, indicating that the scale possesses discriminant validity.

Discussion

In this study, a four-dimensional Physical Education Teachers' Self-Efficacy Scale was developed based on Bandura's social cognitive theory, specifically tailored for physical education teachers. The four factors—teaching process management, student motivation, field expertise, professional development—distinctly

differ from existing general teacher self-efficacy scales (Tschannen-Moran & Hoy, 2001) by comprehensively reflecting the unique contextual elements of physical education teaching, such as physical interaction, safety risks, inclusive practices, facility limitations, and technology integration. This structure concretizes Bandura's principle of reciprocal determinism within the context of physical education and more accurately represents the practical daily challenges faced by teachers.

The exploratory factor analysis indicated that the scale has a four-factor structure, which explains 78.799% of the total variance (Table 4). This proportion is considerably higher than the ranges reported in many scale development studies. In the literature, explained variance between 40% and 60% is generally considered acceptable for a scale (Field, 2024). Similar studies have typically reported variance explained in the range of 50%–60% (Esentürk et al., 2019; Sayan et al., 2022). Although this high variance suggests that the scale strongly represents the self-efficacy construct of physical education teachers, it also raises the possibility of inflated shared variance due to sample characteristics (predominantly male teachers and a majority of middle school participants) or high inter-item correlations. Examination of the sub-dimensions revealed that the extracted factors align well with the self-efficacy dimensions of physical education teachers (Karaiskos et al., 2024). Similar multidimensional structures have been reported in models developed by Pan (2014). Additionally, self-efficacy scales developed within the context of inclusive physical education have also demonstrated a multi-factor structure (Block et al., 2013). These findings indicate that the identified factors make a meaningful contribution to the total variance and are consistent with previous literature.

During the scale development process, factor loadings between 0.30 and 0.45 are generally considered acceptable for establishing a factor structure (Büyüköztürk, 2023). In the present study, a cutoff value of 0.45 was applied, and following rotation, items were clearly associated with their respective factors. The obtained factor loadings substantially exceeded the established threshold, which aligns with previous item analysis findings from the Physical Education Teaching Efficacy Scale and demonstrates consistency with the existing literature (Humphries et al., 2012). Based on these results, a 24-item draft scale was constructed. Specifically, items 1–6 represented the 'Teaching Process Management' dimension, items 7–12 represented 'Student Motivation,' items 13–18 represented 'Field Expertise,' and items 19–24 represented 'Professional Development' (Table 5).

During the confirmatory factor analysis (CFA) of the draft scale, the four-dimensional model comprising 24 items was validated, taking the recommended modification indices into account. Although the initial CFA fit indices were generally acceptable, modification indices were examined to enhance model fit, revealing significant covariance relationships between some items. The resulting fit values are consistent with those reported in international studies on teacher self-efficacy scales (Figure 1, Table 7). Similar CFA fit indices were reported by Woodcock et al. (2025), who emphasized the validity of multidimensional self-efficacy models. Furthermore, achieving the threshold values suggested by Hu and Bentler (1999) and Schermelleh-Engel et al. (2003) indicates that the scale possesses strong structural validity. These findings support the theoretical framework of the scale and confirm that the four-dimensional, 24-item model aligns with the literature.

The internal consistency of the scale was assessed using Cronbach's alpha coefficient, which ranges from 0 to 1, with higher values indicating greater reliability. In the literature, an alpha coefficient of at least 0.70 is generally considered adequate (Büyüköztürk, 2023; Field, 2024). The results obtained in this study are consistent with these standards (Table 6) and indicate that the scale and its sub-dimensions possess both consistent and discriminative properties. These reliability levels align with prior research; for example, Humphries et al. (2012) reported that alpha coefficients above 0.70 are sufficient for scales targeting physical education teachers, and the Interpersonal Self-Efficacy Scale for Physical Education Teachers developed by Esentürk et al. (2019) demonstrated alpha values ranging from 0.78 to 0.87. Accordingly, the reliability results of the present study are in agreement with other scale development studies in the literature.

The sub-dimensions identified in this study and their corresponding labels generally align with previous literature, yet exhibit some notable differences. The findings partially parallel the Physical Education Teaching Efficacy Scale (PETES) developed by Humphries et al. (2012), which addresses instructional, motivational, and Field Expertise dimensions. In contrast, the present scale additionally emphasizes a 'Professional Development' dimension and more explicitly reflects the socio-cultural and institutional realities of Turkey, such as large class sizes, disparities in equipment, and national curriculum dynamics.

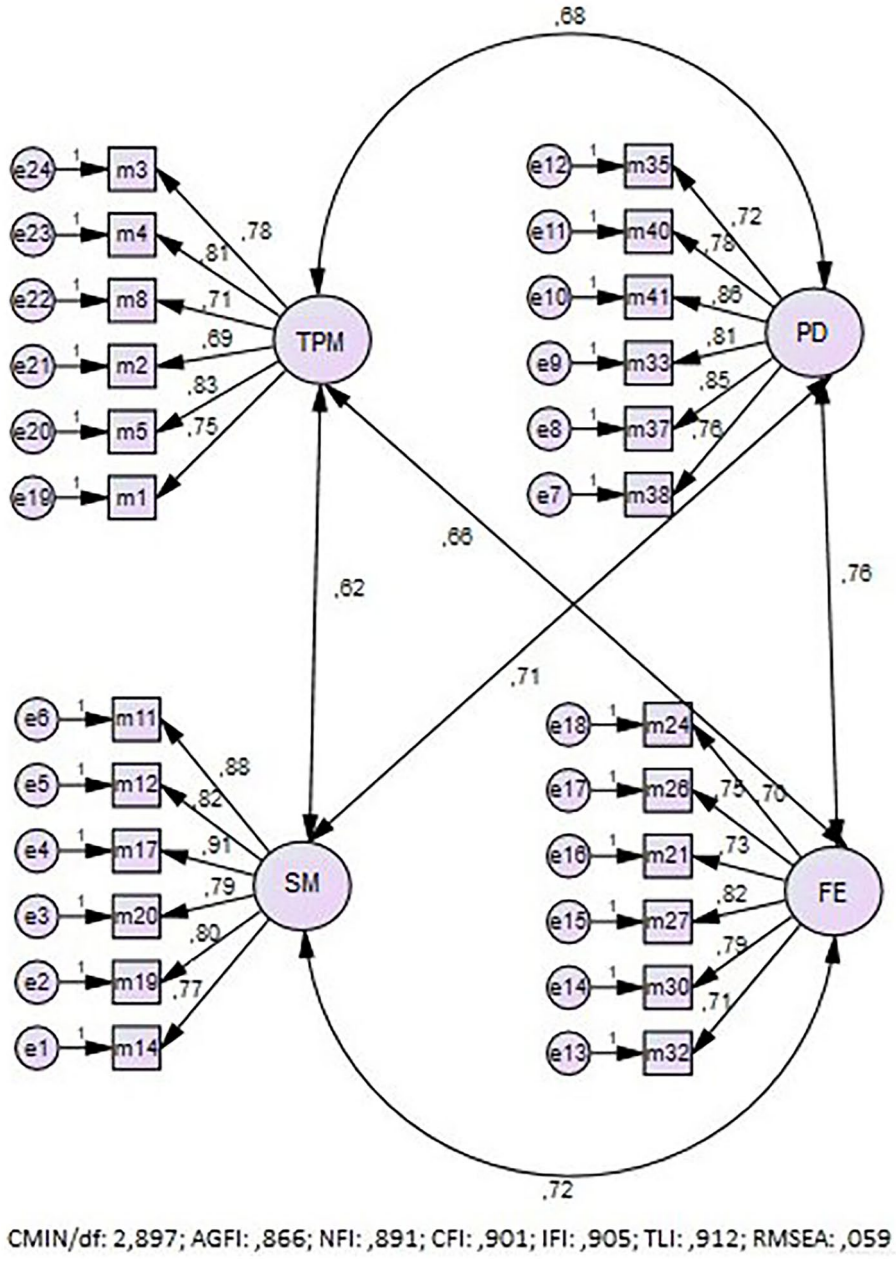

Figure 1. DFA diagram for factor–item relationship.

Table 7. DFA Fit index values.

Fit indices	Scale index values	Excellent fit criteria	Good fit criteria	Result
x^2/df	2.897	<2.5	<3	Good fit
AGFI	0.866	>0.95	>0.85	Good fit
GFI	0.891	>0.95	>0.85	Good fit
NFI	0.901	>0.95	>0.85	Good fit
CFI	0.905	>0.95	>0.85	Good fit
IFI	0.905	>0.95	>0.85	Good fit
TLI	0.912	>0.95	>0.85	Good fit
RMSEA	0.059	<0.05	<0.008	Good fit

This additional dimension is particularly valuable in highlighting the ongoing need for professional learning and adaptive capacity among teachers in developing countries. While national studies, such as the Interpersonal Self-Efficacy Scale developed by Esentürk et al. (2019), tend to focus on more limited dimensions, the current scale encompasses a broader spectrum of professional competencies. Kusmiyati

et al. (2024) reported that physical education teachers' self-efficacy levels directly influence instructional effectiveness and student motivation, supporting the theoretical relevance of the sub-dimensions included in the present scale. Similarly, Yalvuç and Karli (2022) demonstrated that teachers' professional self-efficacy predicts innovative attitudes and instructional quality, indicating that the 'Professional Development' sub-dimension aligns with current educational needs in the field. Recent research on technology integration (Kusmiyati et al., 2024; Pan, 2014) further suggests that teachers' self-efficacy encompasses not only pedagogical Field Expertise but also the ability to support instructional processes through technology. In this respect, the 'Professional Development' dimension highlights, in line with Yalvuç and Karlı, the importance of innovative attitudes and lifelong learning, which may help mitigate professional burnout. Conversely, the 'Teaching Process Management' and 'Field Expertise' dimensions are largely absent in prior literature, yet they capture practice-based challenges specific to physical education, such as physical interaction, equipment usage, and safety concerns. This represents a significant contribution in addressing a gap in the existing literature. However, the high factor loadings observed for these dimensions may partly result from substantial item overlap. This suggests that future studies should consider more comprehensive refinements aimed at increasing item discriminability.

The Physical Education Teachers' Self-Efficacy Scale consists of four sub-dimensions: teaching process management, student motivation, field expertise, professional development. The scale does not include any reverse-coded items; all statements are positively worded, and respondents rate each item using a 5-point Likert-type scale (1 = Strongly Disagree, 5 = Strongly Agree). The scale contains a total of 24 items, with possible total scores ranging from 24 to 120. Each sub-dimension consists of 6 items, resulting in a score range of 6–30 for each subscale. Higher scores on the scale indicate strong perceptions of self-efficacy among physical education teachers in their professional lives, whereas lower scores may suggest potential difficulties in instructional processes, student adaptation, or professional competencies. While the total score reflects overall self-efficacy levels, sub-dimension scores allow for more detailed analysis and interpretation.

Conclusion

From a psychometric perspective, the scale is reliable and theoretically aligned with current literature. Compared to existing instruments with limited structures, it holistically assesses physical education teachers' professional practices, student management, motivational strategies, subject matter expertise, and professional development. In this respect, it fills an important gap in the literature.

Beyond its academic contribution, the scale also has practical implications. By identifying teachers' self-efficacy levels, it enables needs-based planning of in-service training and supports improvements in lesson quality and student engagement through targeted interventions. It is also considered to have the potential to provide scientific evidence for the teacher education, professional development, and performance evaluation processes of the Ministry of National Education. However, to further strengthen its psychometric properties, test–retest reliability studies, advanced statistical analyses, and validation on different samples are recommended. In addition, adapting the scale into different languages and cultural contexts is considered an important research direction that would enhance the generalizability of the findings.

Limitations

This study has several limitations. First, the sample was selected using a non-probability (purposeful) sampling method, which limits the generalizability of the findings to all physical education teachers in Turkey. Although diversity in terms of gender, professional experience, school type, and location was ensured, cross-validation with larger and randomly selected samples is recommended for future studies. Second, the test-retest reliability of the scale was not evaluated; future research should address this to examine the temporal stability of the instrument. Third, data were collected online, which may have introduced social desirability bias. Additional studies assessing convergent and discriminant validity with related constructs-such as teacher burnout, student motivation scales, or general teacher self-efficacy measures-would be beneficial. Finally, the study employed a cross-sectional design; longitudinal research is needed to explore causal relationships.

Acknowledgments

We would like to thank Princess Nourah bint Abdulrahman University for supporting this project through Princess Nourah bint Abdulrahman University Researchers Supporting Project number (PNURSP2026R286), Princess Nourah bint Abdulrahman University, Riyadh, Saudi Arabia.

Institutional review board statement

The study protocol was approved by the Human Research Ethics Committee of Aksaray University Rectorate with the decision dated 18.11.2025 and protocol number 2025-453.

Informed consent statement

Informed consent was obtained from all subjects involved in the study.

Authors contributions

CRediT: **Veli Başal**: Conceptualization, Data curation, Formal analysis, Methodology, Resources, Writing – original draft, Writing – review & editing; **Hüseyin Ünlü**: Conceptualization, Data curation, Formal analysis, Investigation, Project administration, Supervision; **Metin Yüceant**: Formal analysis, Methodology, Software, Validation, Visualization; **Monira I. Aldhahi**: Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Supervision, Validation.

Disclosure statement

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

Funding

This research was funded by the Princess Nourah bint Abdulrahman University Researchers Supporting Project number (PNURSP2026R286), Princess Nourah bint Abdulrahman University, Riyadh, Saudi Arabia.

ORCID

Veli Başal http://orcid.org/0000-0003-3536-0500
Hüseyin Ünlü http://orcid.org/0000-0002-3079-4417
Metin Yüceant http://orcid.org/0000-0002-6867-2183
Monira I. Aldhahi http://orcid.org/0000-0002-5255-4860

Data availability statement

All the data obtained and produced in the scope of this study has been not deposited into a publicly available repository and the data will be made available on request.

References

Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review*, *84*(2), 191–215. https://doi.org/10.1037/0033-295x.84.2.191

Bandura, A. (1997). *Self-efficacy: The exercise of control* (Vol. 11). Freeman.

Bandura, A. (2006). Guide for constructing self-efficacy scales. *Self-Efficacy Beliefs of Adolescents*, *5*(1), 307–337.

Beaton, D. E., Bombardier, C., Guillemin, F., & Ferraz, M. (2000). Guidelines for the process of cross-cultural adaptation of self-report measures. *Spine*, *25*(24), 3186–3191. https://doi.org/10.1097/00007632-200012150-00014

Block, M. E., Hutzler, Y., Barak, S., & Klavina, A. (2013). Creation and validation of the self-efficacy instrument for physical education teacher education majors toward inclusion. *Adapted Physical Activity Quarterly*, *30*(2), 184–205. https://doi.org/10.1123/apaq.30.2.184

Braksiek, M. (2022). Pre-service physical education teachers' attitude toward, and self-efficacy in, inclusive physical education: Measurement invariance and influence factors. *Teaching and Teacher Education*, *109*, 103547. https://doi.org/10.1016/j.tate.2021.103547

Büyüköztürk, Ş. (2023). *Sosyal bilimler için veri analizi el kitabı*. Pegem Atıf İndeksi.

Büyüköztürk, Ş., Akgün, Ö. E., Demirel, F., Karadeniz, Ş., & Çakmak, E. K. (2022). *Bilimsel araştırma yöntemleri* [Scientific research methods]. Pegem Akademi.

Ercan, İ., & Kan, İ. (2004). Ölçeklerde güvenirlik ve geçerlik. *Uludağ Üniversitesi Tıp Fakültesi Dergisi*, *30*(3), 211–216.

Eroğlu, C., & Ünlü, H. (2015). Self-efficacy: Its effects on physical education teacher candidates' attitudes toward the teaching profession. Educational Sciences, 15(1), 201–212.

Esentürk, O. K., Yılmaz, A., İlhan, E. L., & Kan, A. (2019). Beden eğitimi öğretmeni kişilerarası öz-yeterlik ölçeği'nin psikometrik özelliklerinin incelenmesi. *Gazi Journal of Physical Education and Sport Sciences*, *24*(2), 63–88.

Field, A. (2024). *Discovering statistics using IBM SPSS statistics*. Sage Publications Limited.

Fornell, C., & Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research*, *18*(1), 39–50. https://doi.org/10.1177/002224378101800104

Foster, J. J. (2001). *Data analysis using SPSS for Windows versions 8–10: A beginner's guide.* SAGE Publications.

Guskey, T. R., & Passaro, P. D. (1994). Teacher efficacy: A study of construct dimensions. *American Educational Research Journal*, *31*(3), 627–643. https://doi.org/10.3102/00028312031003627

Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, *6*(1), 1–55. https://doi.org/10.1080/10705519909540118

Humphries, C. A., Hebert, E., Daigle, K., & Martin, J. (2012). Development of a physical education teaching efficacy scale. *Measurement in Physical Education and Exercise Science*, *16*(4), 284–299. https://doi.org/10.1080/1091367X.2012.716726

Karaiskos, L., Sotiras, M. N., Antonopoulou, P., & Gdonteli, K. (2024). The impact of training and professional development on physical education teacher self-efficacy. *Journal of Physical Education and Sport*, *24*(11), 2022–2033.

Karasar, N. (2009). *Bilimsel araştırma yöntemi: Kavramlar-ilkeler-teknikler*. Nobel Yayın Dağıtım.

Kartal, S. K., & Dirlik, E. M. (2016). Geçerlik kavramının tarihsel gelişimi ve güvenirlikte en çok tercih edilen yöntem: Cronbach alfa katsayısı. *Abant İzzet Baysal Üniversitesi Eğitim Fakültesi Dergisi*, *16*(4), 1865–1879.

Kusmiyati, K., Setiawati, H., Setijono, H., Soedjatmiko, S., Kristiyanto, A., & Suryadi, D. (2024). Self-efficacy of physical education teachers: How does it contribute to achieving the overarching objectives of national sport? *Retos*, *58*, 546–551. https://doi.org/10.47197/retos.v58.108159

Lawshe, C. H. (1975). A quantitative approach to content validity. *Personnel Psychology*, *28*(4), 563–575. https://doi.org/10.1111/j.1744-6570.1975.tb01393.x

Lynn, M. R. (1986). Determination and quantification of content validity. *Nursing Research*, *35*(6), 382–385. https://doi.org/10.1097/00006199-198611000-00017

Pan, Y. H. (2014). Relationships among teachers' self-efficacy and students' motivation, atmosphere, and satisfaction in physical education. *Journal of Teaching in Physical Education*, *33*(1), 68–92. https://doi.org/10.1123/jtpe.2013-0069

Polit, D. F., & Beck, C. T. (2006). The content validity index: are you sure you know what's being reported? Critique and recommendations. *Research in Nursing & Health*, *29*(5), 489–497. https://doi.org/10.1002/nur.20147

Sayan, S., Akay, B., Ceylan, M., & Orhan, R. (2022). Beden eğitimi öğretmeni adayları için özel öğretim yöntemleri öz-yeterlik ölçeği: Geliştirme ve güvenirlik çalışması. *Uluslararası Güncel Eğitim Araştırmaları Dergisi*, *8*(2), 624–635.

Schermelleh-Engel, K., Moosbrugger, H., & Müller, H. (2003). Evaluating the fit of structural equation models: Tests of significance and descriptive goodness-of-fit measures. *Methods of Psychological Research Online*, *8*(2), 23–74.

Şeker, H., Deniz, S., & Görgen, İ. (2004). Öğretmen yeterlikleri ölçeği. *Milli Eğitim Dergisi*, *164*, 105–118.

Tabachnick, B. G., & Fidell, L. S. (2013). *Using multivariate statistics (6. Baskı)*. Pearson.

Tavşancıl, E. (2014). *Tutumların ölçülmesi ve SPSS ile veri analizi (5. baskı)*. Nobel Yayınevi.

Tschannen-Moran, M., & Hoy, A. W. (2001). Teacher efficacy: Capturing an elusive construct. *Teaching and Teacher Education*, *17*(7), 783–805. https://doi.org/10.1016/S0742-051X(01)00036-1

Ünlü, H., Sünbül, A. M., & Aydos, L. (2008). Beden eğitimi öğretmenleri yeterlilik ölçeği: Geçerlilik ve güvenirlik çalışması. *Ahi Evran Üniversitesi Kırşehir Eğitim Fakültesi Dergisi*, *9*(2), 23–33.

Woodcock, S., Tournaki, N., & Ehrich, J. (2025). Teacher self-efficacy: Validating a new measurement scale to capture the elusive construct. *British Educational Research Journal*, *51*(5), 2229–2255. https://doi.org/10.1002/berj.4163

Yalvuç, F., & Karli, Ü. (2022). Beden eğitimi ve spor öğretmenlerinde mesleki öz-yeterlik ve yenilikçi tutum. *Sportif Bakış: Spor ve Eğitim Bilimleri Dergisi*, *9*(3), 294–312. https://doi.org/10.33468/sbsebd.283

Yenen, E. T. (2022). Öğretmenlerin mesleki yeterliklerini etkileyen faktörler. *MANAS Sosyal Araştırmalar Dergisi*, *11*(1), 27–45. https://doi.org/10.33206/mjss.767720

Zee, M., & Koomen, H. M. (2016). Teacher self-efficacy and its effects on classroom processes, student academic adjustment, and teacher well-being: A synthesis of 40 years of research. *Review of Educational Research*, *86*(4), 981–1015. https://doi.org/10.3102/0034654315626801