Ahi Evran Üniversitesi Sosyal Bilimler Enstitüsü Dergisi (AEÜSBED)
2026, Cilt 12, Sayı 2, Sayfa 469-487

Development and Validation of the Self-Regulated Teaching Scale (SRTS)

İhsan AKEREN
Bayburt Üniversitesi
aydostdeyince@gmail.com
ORCID ID: 0000-0001-5615-4189

Araştırma Makalesi		DOI: 10.31592/aeusbed.1754726
Geliş Tarihi: 31.07.2025	Revize Tarihi: 13.03.2026	Kabul Tarihi: 19.05.2026

Atıf Bilgisi
Akeren, İ. (2026). Development and validation of the self-regulated teaching scale (SRTS). *Ahi Evran Üniversitesi Sosyal Bilimler Enstitüsü Dergisi, 12*(2), 469-487.

ABSTRACT

Quality education depends on the interaction among the student, the teacher, and the curriculum. Within this process, teachers play a central role in planning instruction, implementing teaching, and evaluating learning. In this context, self-regulated teaching refers to the deliberate regulation of teaching practices across the planning, implementation, and evaluation phases. This study aimed to develop and validate a scale measuring teachers' self-regulated teaching skills on the basis of Zimmerman's Self-Regulated Learning Model (Z'SRLM). An exploratory sequential mixed-methods design was employed. Qualitative data from three focus group discussions with 12 teachers were used to generate the item pool, yielding three dimensions: planning, implementation, and evaluation. After expert review and a clarity check with 75 teachers, exploratory factor analysis (EFA) was conducted with 226 teachers and produced a 17-item, three-factor structure. Confirmatory factor analysis (CFA) with an independent sample of 394 teachers supported the final modified three-factor model (χ^2/df = 1.374, CFI = .962, TLI = .955, IFI = .963, RMSEA = .031, SRMR = .038). The correlated three-factor and second-order representations yielded equivalent global fit, and additional model-based indices indicated that the total score retained meaningful psychometric value as a complementary global indicator, although the scale was interpreted primarily at the subdimension level. Criterion-related validity was examined with 207 teachers using the Classroom Management Skills Scale (CMSS), and the correlations were positive and statistically significant (r = .438-.751, p < .001). Internal consistency coefficients ranged from acceptable to high across subdimensions, model-based indices supported cautious interpretation of the total score, and test-retest findings indicated acceptable temporal stability. Overall, the findings support the validity and reliability of the Self-Regulated Teaching Scale for use in research on teachers' instructional practices.
Keywords: Self-regulation, teaching, self-regulated teaching, validity, reliability.

Öz Düzenlemeli Öğretim Becerileri Ölçeği (ÖDÖBÖ): Geliştirme ve Geçerlilik Çalışması

ÖZ

Kaliteli eğitim, öğrenci, öğretmen ve programın birbiriyle uyumlu etkileşimine bağlıdır. Bu süreçte öğretmen, öğretimi planlama, uygulama ve değerlendirme boyutlarında merkezi bir rol üstlenmektedir. Bu bağlamda öz düzenlemeli öğretim, öğretim uygulamalarının planlama, uygulama ve değerlendirme aşamalarında bilinçli biçimde düzenlenmesini ifade etmektedir. Bu çalışmada, Zimmerman'ın Öz Düzenlemeli Öğrenme Modeli'ne (Z'SRLM) dayalı olarak öğretmenlerin öz düzenlemeli öğretim becerilerini ölçen bir ölçek geliştirilmesi ve ölçeğin geçerlik ve güvenirlik kanıtlarının incelenmesi amaçlanmıştır. Araştırmada keşfedici sıralı karma yöntem deseni kullanılmıştır. İlk aşamada, 12 öğretmenle gerçekleştirilen üç odak grup görüşmesinden elde edilen nitel veriler doğrultusunda madde havuzu oluşturulmuş ve planlama, uygulama ve değerlendirme olmak üzere üç boyut belirlenmiştir. Uzman görüşü ve 75 öğretmenle gerçekleştirilen açıklık-anlaşılırlık incelemesinden sonra, 226 öğretmenden elde edilen verilerle açımlayıcı faktör analizi yapılmış ve 17 maddelik üç faktörlü yapı elde edilmiştir. Bağımsız bir 394 öğretmen örnekleminde gerçekleştirilen doğrulayıcı faktör analizi, nihai modifikasyonlu üç faktörlü modeli desteklemiştir (χ^2/sd = 1.374, CFI = .962, TLI = .955, IFI = .963, RMSEA = .031, SRMR = .038). Korelasyonel üç faktörlü model ile ikinci düzey modelin eşdeğer uyum verdiği görülmüş; model temelli ek göstergeler toplam puanın alt boyut puanlarına göre ikincil ve tamamlayıcı bir genel gösterge olarak psikometrik açıdan anlamlı bir temsil sunduğunu göstermiştir. Ölçüt ilişkili geçerlik, 207 öğretmenden elde edilen verilerle Sınıf Yönetimi Becerileri Ölçeği (SYBÖ) kullanılarak incelenmiş; korelasyonların pozitif ve anlamlı olduğu görülmüştür (r = .438-.751, p < .001). İç tutarlılık katsayıları alt boyutlar düzeyinde kabul edilebilir ile yüksek düzey arasında değişmiş, model temelli ek göstergeler toplam puanın dikkatli biçimde yorumlanabileceğini göstermiş ve test-tekrar test bulguları kabul edilebilir zamansal kararlılık ortaya koymuştur. Bulgular, Öz Düzenlemeli Öğretim Becerileri Ölçeği'nin geçerli ve güvenilir bir ölçme aracı olduğunu göstermektedir.
Anahtar Kelimeler: Öz düzenleme, öğretim, öz düzenlemeli öğretim, geçerlilik, güvenilirlik.

Introduction

Teaching is a complex, student-oriented process shaped by the interaction among content, method, environmental conditions, and the teacher (Büyükkaragöz ve Çivi, 1999). Its basic elements are the student, the teacher, and the curriculum, and effective teaching depends on the compatibility and cooperation of these elements (Demirel, 2009). Although teaching may sometimes be viewed as the transmission of knowledge, contemporary teaching involves much more than information transfer and requires teachers to regulate multiple instructional processes (Ergün, Ergezer, Çevik ve Özdaş, 1999).

Ensuring effectiveness in teaching and learning is a central aim of education systems, and research on teaching effectiveness generally places teachers at the center of this process (Cothran and Kulinna, 2008; Lesh and Lehrer, 2003). Teachers organize learning environments, plan and implement instruction, support students' development, and evaluate instructional effectiveness (Karakelle, 2005). Studies on teacher quality have drawn on evidence from teachers, students, and parents (Liu and Meng, 2009), classroom observations (Strong, Gargani and Hacifazlioğlu, 2011), and theoretical syntheses (Stronge, 2018). Across these perspectives, teacher quality has been associated with personality traits, relationship skills, problem-solving, flexibility, and instructional competence (Banks and Thompson, 1995; Demirel, 2009). The present study focuses specifically on the cognitive-regulatory dimension of teaching, including planning instruction, adapting strategies to contextual and student-related conditions, perceiving oneself as competent, remaining flexible, and evaluating both student progress and one's own teaching. These features are closely aligned with self-regulation.

Self-regulation

Self-regulation has been widely examined since Bandura (1986) defined it within a social-cognitive learning perspective. According to Bandura, self-regulation is an internal system through which individuals manage their thoughts, feelings, and behaviors and learn from their environment. Over time, several self-regulation-based learning models have been developed, including the models proposed by Boekaerts (1997), Borkowski (1996), and Pintrich (2000). Among these, Zimmerman's Self-Regulated Learning Model (Z'SRLM) is particularly influential. Z'SRLM consists of three cyclical phases: forethought, performance, and self-reflection (Zimmerman, 2000; Zimmerman and Campillo, 2003). These phases include task analysis and self-motivation, self-control and self-observation, and self-judgement and self-reaction. The cyclical nature of the model means that self-reflection affects the next forethought phase (Zimmerman, 2013).

Research shows that students' self-regulation skills are positively related to academic achievement and negatively related to test anxiety (El-Adl and Alkharusi, 2020). Similar findings indicate that self-regulated learning strategies improve perceived achievement (Zheng, Ward and Stanulis, 2020) and that self-regulation interventions support academic achievement (Lee, Lee, Kim and Lee, 2023). These findings suggest that teachers should also possess self-regulatory skills. Teachers serve as an important bridge between students and instructional content, and teaching strategies directly shape how students learn (Donche, De Maeyer, Coertjens, Van Daal and Van Petegem, 2013). When pre-service teachers learn self-regulation, they may transfer this skill to their own teaching activities (Randi, 2004). Self-regulated teachers adjust their teaching strategies, encourage students' self-regulated learning, and adopt innovative knowledge and strategies (Peeters, De Backer, Reina, Kindekens, Buffel and Lombaerts, 2014). Prior studies also suggest that teachers' self-regulation is related to classroom management and emotion regulation, which supports viewing it as a professional competence in teaching (Kurt ve Sığırtmaç, 2021; Sutton, Mudrey-Camino and Knight, 2009). Thus, teachers' self-regulation may support both their instructional effectiveness and students' learning outcomes (Filice, Tregunno, Edge and Egan, 2020).

Current Study

Teaching aims to promote student learning and achievement, and many factors related to the school, teacher, family, and student affect this process. The present study focuses on the teacher factor and seeks to develop a measurement tool for assessing teaching practices that may contribute to student success. As teacher quality is an important indicator of student success, the effects of good teachers may continue even after students move to the next grade (Stronge, 2018). Teachers act not merely as transmitters of information but as facilitators who organize learning environments, interact with students, manage instruction, and evaluate outcomes (Kılıç, 2006; Sönmez, 2020). Recent studies also show the continuing need to evaluate teaching programs, curricula, and instructional processes from different perspectives (Aslanderen, Çilsalar Sagnak ve Sağlam, 2024; Güneş ve Demir, 2025).

Teachers engage in planning and preparation before instruction (Christie, 2014), use various strategies during teaching (Grant, Hindman and Stronge, 2013), and evaluate instructional outcomes afterwards through examinations, questioning, and observation (Tuckman, 1995). These recurring phases can be conceptualized as planning, implementation, and evaluation. Because they closely overlap with the cyclical logic of Z'SRLM, they provide a suitable framework for conceptualizing self-regulated teaching. Despite the growing literature on self-regulated learning and teacher effectiveness, no concise instrument directly aligned with Z'SRLM was identified in the literature reviewed for this study to assess teachers' self-regulated teaching practices. Existing instruments mostly address broad teacher competencies, classroom management, or general self-regulation rather than teachers' deliberate regulation of instructional practice across the planning, implementation, and evaluation phases. Therefore, this study aimed to examine whether teachers' instructional practices could be conceptually organized in line with Z'SRLM, test whether these practices formed a valid latent structure, and develop a psychometrically sound instrument for assessing self-regulated teaching.

Method

The method section presents the research design, validity procedures, study groups, data collection tools, data analysis process, and ethical considerations. Because this study aimed to develop and validate a scale, qualitative and quantitative procedures are reported together to show how the item pool and psychometric evidence were generated. The following subsections describe these procedures in detail.

Research Model and Design

An exploratory sequential mixed-methods design was employed because the study first required qualitative evidence for item generation and then quantitative evidence for scale validation (Creswell, 2017). In the first phase, qualitative data were obtained through three focus group discussions with 12 teachers and were analyzed to generate themes, codes, and an initial item pool. In the second phase, the resulting items were subjected to expert review, a preliminary clarity check, exploratory factor analysis, confirmatory factor analysis, and criterion-related validity and reliability analyses. Because measurement in the social sciences is theory-driven, the construct of self-regulated teaching was conceptualized on the basis of Zimmerman's Self-Regulated Learning Model. The procedures followed to establish content, construct, and criterion-related validity, as well as reliability, are presented below.

Validity

In the present study, validity evidence was considered in terms of content, construct, and criterion-related validity, in line with common psychometric classifications (DeVellis, 2017; Garson, 2013). The scale development process was also guided by recommendations concerning item generation, questionnaire design, expert review, and reporting standards in scale development research

(Carpenter, 2018; Rattray and Jones, 2007; Seçer, 2018; Singh, Junnarkar and Kaur, 2016; Worthington and Whittaker, 2006).

Focus Group Discussions for Content Validity

Content validity refers to the extent to which scale items adequately represent the construct and its conceptual domain (DeVellis, 2017). To generate the initial item pool, three focus group discussions were conducted with 12 in-service teachers: one group of primary school teachers (n = 3), one group of middle school teachers (n = 5), and one group of high school teachers (n = 4). Each session was held in the teachers' room, lasted about 60 minutes, and was guided by three prompts focusing on what teachers do before, during, and after instruction. All discussions were moderated by the researcher using the same semi-structured guide, and participants were encouraged to interact with one another. The participants represented different subject areas, were between 26 and 52 years old, and had between 3 and 31 years of teaching experience.

The qualitative data were coded manually by three independent coders, following common procedures for qualitative data analysis (Miles, Huberman and Saldaña, 2014). The data were first reviewed individually, then organized into codes, and finally grouped under broader themes related to pre-instruction, during-instruction, and post-instruction practices. Since the interview questions were already structured around these three phases, there was full agreement at the theme level. At the code level, the coders showed a high degree of agreement, and the few differences mostly reflected wording rather than meaning. For instance, some coders used labels such as preparation, teaching, and review, whereas the final labels were standardized as planning, implementation, and evaluation in line with the conceptual framework of the study. No formal consensus meeting was held because the coders were in different cities. Instead, the researcher coordinated the process by e-mail, compared the coding forms, standardized the labels, and shared the finalized version with the coders before constructing the item pool. Inter-coder agreement was calculated as .92.

The themes and codes derived from the focus group discussions were then compared with the cyclical structure of Z'SRLM. This comparison provided the theoretical basis for the initial structure of the scale and supported the alignment of the item pool with the planning, implementation, and evaluation dimensions shown in Figure 1.

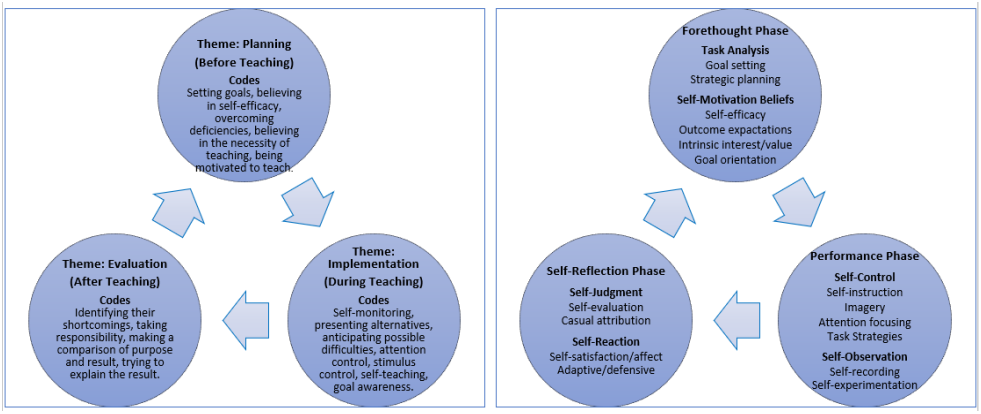


Figure 1. Themes Derived from the Focus Group Discussions and Their Correspondence to Z'SRLM

Expert and Practitioner Opinion for Content Validity

Following the focus group analysis, the emerging themes and codes were reviewed in relation to Zimmerman's Self-Regulated Learning Model to support the content coverage of the scale. On this basis, an initial pool of 32 items was drafted in a five-point Likert format. The wording of the items was first reviewed by a language expert for clarity and accuracy. The revised items were then evaluated by three scholars with doctoral degrees in teacher education using Davis's (1992) criteria.

Based on these ratings, two items with a content validity coefficient below .80 were removed. The resulting 30-item draft form was subsequently reviewed by 75 teachers to evaluate clarity, comprehensibility, and wording. This stage was intended to improve item expression rather than to perform statistical item screening; therefore, item-total correlations were not calculated at this stage. Based on participant feedback, several minor wording revisions were made, but no additional items were removed.

Study Group

This exploratory sequential scale development study was conducted with five independent samples. The first sample included 12 in-service teachers who participated in three focus group discussions for item generation. The second sample consisted of 75 teachers who reviewed the preliminary form for clarity, comprehensibility, and wording. The third sample included 226 teachers (111 female, 115 male; age range = 20-59 years, M = 39.80, SD = 7.97) and was used for exploratory factor analysis. The fourth sample included 394 teachers (201 female, 193 male; age range = 23-64 years, M = 42.99, SD = 12.39) and was used for confirmatory factor analysis and internal consistency analyses. The fifth sample included 207 teachers (106 female, 101 male; age range = 21-61 years, M = 39.08, SD = 11.81) and was used for criterion-related validity analyses. A subsample of 93 teachers from this group (48 female, 45 male; age range = 21-61 years, M = 39.81, SD = 12.14) completed the scale again after a 4-6-week interval for test-retest analysis.

Participation was voluntary in all phases, and all participants were informed about the purpose of the study and ethical approval. The inclusion criterion was active employment in a teaching role at the primary, middle, or high school level in public schools, private schools, or private teaching institutions. Preschool and kindergarten teachers, school psychological counsellors, and educators working outside these school levels were excluded. School administrators were included only if they were actively engaged in teaching duties. To strengthen representativeness, teachers from 61 cities in Türkiye were reached during the exploratory factor analysis phase through cluster sampling. Given the 30-item draft form, the exploratory factor analysis sample of 226 teachers exceeded commonly recommended minimums for factor-analytic procedures in scale development studies (Beavers, Lounsbury, Richards, Huck, Skolits and Esquivel, 2013; MacCallum, Widaman, Zhang and Hong, 1999). More detailed contextual information such as school level and settlement type was available for the exploratory factor analysis sample, whereas only age and gender were collected for the later quantitative subsamples to reduce respondent burden.

Classroom Management Skills Scale (CMSS)

The scale developed by Delson-Karan (1984) and adapted into Turkish by Yalçınkaya ve Tonbul (2002) has six factors (physical conditions: PHC, efficient use of time: EUT, presentation: PRS, attendance in class: ATC, asking questions: ASQ and psychological factors: PSF) and consists of 25 items. The items are scored on a 4-point Likert scale from 0 to 4. The lowest score is 0 and the highest score is 100. High scores mean that teachers are successful in classroom management. In the present study, the CMSS items were administered as a self-report, with teachers asked to rate their own classroom management behaviors as if observing them from an external perspective. In the adaptation study, the Cronbach's Alpha coefficient of the scale was .88 (Yalçınkaya ve Tonbul, 2002). In this study, reliability coefficients were PHC = .65, EUT = .73, PRS = .78, ATC = .66, ASQ = .72, PSF = .72, and overall α = .94, indicating acceptable internal consistency.

Data Collection and Analysis

Quantitative analyses were conducted using IBM SPSS Statistics 27, and confirmatory factor analyses and comparative model testing were conducted using JASP (Version 0.95.4). Prior to the main analyses, the data were examined in terms of descriptive statistics, normality, and suitability for factor analysis. Multivariate screening and factor-analytic decisions were guided by common recommendations for applied multivariate analysis and exploratory factor analysis (Watkins, 2021).

The suitability of the exploratory factor analysis data was evaluated using the Kaiser-Meyer-Olkin measure of sampling adequacy and Bartlett's test of sphericity.

Following expert review, two items were removed from the initial 32-item pool, resulting in a 30-item draft form. After the clarity and comprehensibility review with 75 teachers, exploratory factor analysis was conducted using principal axis factoring with Direct Oblimin rotation because the dimensions of self-regulated teaching were theoretically expected to be correlated. Items were removed iteratively if they had factor loadings below .32 or problematic cross-loadings, defined as a difference smaller than .10 between the two highest loadings. Factor retention was evaluated using the scree plot, parallel analysis, theoretical interpretability, and the overall pattern of factor loadings, consistent with recommendations for scale development and factor-analytic practice (Beavers, Lounsbury, Richards, Huck, Skolits and Esquivel, 2013; Watkins, 2021). The final decision was based on the convergence of statistical evidence, simple structure, theoretical coherence, and later CFA comparisons among alternative models.

Because the items were ordinal, CFA was conducted using the weighted least squares mean and variance adjusted estimator (WLSMV). For gender invariance, robust maximum likelihood estimation with full information maximum likelihood (FIML) was preferred because it produced stable multigroup solutions in JASP. CFA was conducted on an independent sample to test the three-factor structure obtained in the exploratory phase. The unmodified three-factor model, modified three-factor model, unmodified single-factor model, and modified single-factor model were tested and compared using χ^2/df, CFI, TLI, IFI, RMSEA, and SRMR values. The interpretation of CFA results and model fit indices was based on commonly used criteria in structural equation modeling and confirmatory factor analysis (Brown, 2015; Hu and Bentler, 1999; Kline, 2023).

Because the correlated three-factor CFA model and the second-order CFA model yielded empirically equivalent global fit, additional model-based analyses were conducted to evaluate the interpretability of the total score. Specifically, McDonald's omega total, omega hierarchical, explained common variance (ECV), and percentage of uncontaminated correlations (PUC) were used to evaluate the psychometric meaning of the total score beyond the three specific dimensions (Rodriguez, Reise and Haviland, 2016a, 2016b). Convergent and discriminant validity were examined using average variance extracted (AVE), composite reliability (CR), and the heterotrait-monotrait ratio of correlations (HTMT).

Because both the SRTS and the CMSS were administered as self-report measures, common method variance was checked using Harman's single-factor test, and the first unrotated factor accounted for 24.45% of the total variance. Measurement invariance across gender was examined through multigroup CFA by testing configural, metric, and scalar models. Criterion-related validity was examined through correlation and multiple regression analyses with the CMSS. Before regression analysis, multicollinearity, residual normality, independence of errors, linearity, and homoscedasticity assumptions were examined.

Research Ethics

Ethical approval for this study was obtained from the Bayburt University Ethics Committee (09.06.2023, Decision No. 219-10, No. 137345). The study was conducted in accordance with the Declaration of Helsinki and subsequent relevant ethical principles. Participants were informed about the purpose of the study, participation was voluntary, and no identifying personal information was collected.

Findings

Exploratory Factor Analysis (EFA)

Table 1

Factor Loadings and Communalities for the Final 17-Item Scale

Item no (Code)	Item Content	Factor Loadings			
		Planning	Implementation	Evaluation	h^2
1 (i1)	I prepare my plans according to the curriculum (annual-monthly-daily).	.52			.25
2 (i2)	I consider the developmental characteristics of my students while preparing my plans.	.71			.60
3 (i3)	I eliminate my deficiencies, if any, before the teaching activity.	.68			.61
4 (i4)	I feel competent in teaching activities.		.41		.31
5 (i5)	I can teach the objectives determined in the plans to the students.		.72		.52
6 (i6)	I can provide desired behavior change in students with my teaching service.		.58		.46
7 (i8)	I am willing to help students attain the instructional objectives.		.91		.70
8 (i9)	I care about the acquisition of teaching objectives.		.86		.62
9 (i10)	I teach the teaching objectives because I believe in their necessity.		.45		.40
10 (i13)	I give the teaching content in the appropriate order in my teaching.		.52		.46
11 (i14)	I am aware of the objectives I set in the teaching process.		.56		.66
12 (i15)	I conduct my teaching process in line with the objectives I set.		.55		.58
13 (i23)	After teaching, I try to determine to what extent my behaviors contribute to my goals.			.53	.60
14 (i24)	I realize my deficiencies, if any, in the teaching process.			.50	.61
15 (i25)	I question whether my teaching activities achieve their aims or not.			.50	.65
16 (i26)	I hold myself responsible for the outcome of the teaching activity when necessary.			.74	.49
17 (i27)	If my students did not learn enough, I question my deficiencies.			.74	.59
	Explained variance (%)/eigenvalue	46.05 / 7.83	4.36 / .74	3.09 / .53	

Note. Exploratory factor analysis was conducted on the data obtained from 226 teachers. No missing responses were observed in the scale items of this subsample; therefore, all cases were retained in the analysis. The Kaiser-Meyer-Olkin measure of sampling adequacy was .93, and Bartlett's test of sphericity was significant, $\chi^2(136) = 2140.45$, $p < .001$, indicating that the data were suitable for factor analysis.

As an additional factor-retention check, FA-based parallel analysis was conducted on the 30-item draft form in JASP. Parallel analysis suggested a five-factor solution. However, when these statistically suggested solutions were examined using the same loading criterion (.32) and cross-loading rule (difference < .10), they yielded fragmented and psychometrically weak structures, including one-item and two-item factors, and did not align with the theoretically expected planning, implementation, and evaluation dimensions derived from Z'SRLM. Therefore, the final retained model was determined on the basis of the joint consideration of scree plot, parallel analysis, item performance, theoretical interpretability, simple structure, and later CFA comparisons among alternative models.

Using principal axis factoring with Direct Oblimin rotation, a three-factor structure consisting of 17 items was obtained as the most interpretable and theoretically coherent final solution. Items with factor loadings below .32 or problematic cross-loadings were excluded during the iterative analysis process. The final solution explained 53.50% of the total variance, with the Planning, Implementation, and Evaluation factors accounting for 46.05%, 4.36%, and 3.09% of the variance, respectively. Although the statistically suggested alternatives included more fragmented solutions, the scree plot, theoretical interpretability, simple structure, and later CFA comparisons supported the retention of the three-factor model. The factor loadings are presented in Table 1. The communalities of the retained items ranged from .25 to .70, indicating that the items shared an acceptable proportion of variance with the extracted factor solution.

Table 2
Descriptive Statistics and Inter-Item Correlations for the Final Scale Items

	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**
1. i1	1																
2. i2	.384	1															
3. i3	.423	.597	1														
4. i4	.168	.355	.412	1													
5. i5	.153	.471	.447	.466	1												
6. i6	.206	.471	.434	.455	.532	1											
7. i7	.225	.416	.476	.465	.558	.552	1										
8. i8	.254	.421	.423	.354	.545	.493	.740	1									
9. i9	.284	.389	.376	.320	.353	.466	.487	.551	1								
10. i10	.265	.368	.396	.354	.468	.400	.487	.474	.462	1							
11. i11	.294	.520	.549	.399	.539	.497	.624	.572	.493	.667	1						
12. i12	.280	.485	.439	.394	.514	.518	.544	.542	.457	.590	.722	1					
13. i13	.273	.457	.507	.337	.439	.454	.471	.415	.453	.505	.562	.574	1				
14. i14	.283	.458	.493	.404	.439	.439	.573	.505	.416	.473	.571	.572	.705	1			
15. i15	.294	.536	.531	.420	.461	.386	.542	.503	.439	.494	.643	.543	.672	.685	1		
16. i16	.177	.280	.361	.365	.234	.361	.340	.377	.366	.312	.398	.363	.443	.463	.483	1	
17. i17	.251	.419	.414	.327	.336	.416	.376	.358	.440	.389	.504	.451	.528	.523	.553	.613	1
M	4.08	4.25	4.38	4.29	4.35	4.22	4.42	4.44	4.21	4.13	4.38	4.37	4.23	4.31	4.42	4.15	4.24
SD	.86	.77	.65	.64	.60	.66	.61	.65	.77	.77	.64	.61	.64	.63	.62	.71	.70
Skewness	-1.25	-1.11	-1.05	-.55	-.44	-.45	-.65	-1.05	-.91	-1.05	-.94	-.66	-.45	-.57	-.81	-.52	-.75
Kurtosis	2.58	1.96	2.65	.34	.08	.13	.13	1.23	1.08	1.76	1.70	.69	.30	.48	.77	.09	1.30

Note. N = 226. Pearson correlation coefficients are presented below the diagonal. M = mean; SD = standard deviation.

Table 2 presents descriptive statistics and inter-item correlations for the final 17 items rather than the factor structure itself. The inter-item correlations were generally moderate to strong, and the skewness and kurtosis values fell within acceptable ranges, indicating no substantial deviation from normality. Conceptually, the first factor represented teachers' planning before instruction, the second factor reflected their implementation of instruction, and the third factor represented their evaluation of teaching after instruction. This three-phase structure was consistent with the cyclical logic of Zimmerman's Self-Regulated Learning Model. Accordingly, the factors were named Planning, Implementation, and Evaluation.

Although Item 1 showed a relatively lower corrected item-total correlation than the remaining items, it was retained because it represented an essential aspect of the planning phase within the theoretical framework of the scale. Item-level reliability evidence is presented in the Reliability section.

Confirmatory Factor Analysis (CFA)

Confirmatory factor analysis was conducted on an independent sample of 394 teachers to test the factor structure obtained in the exploratory phase. The unmodified three-factor model showed acceptable but improvable fit to the data, $\chi^2(116) = 218.700$, $p < .001$, CFI = .908, TLI = .893, IFI = .910, RMSEA = .047, 90% CI [.038, .057], and SRMR = .048. All standardized factor loadings were

statistically significant and ranged from .45 to .84 for Planning, .52 to .82 for Implementation, and .57 to .85 for Evaluation.

Inspection of the modification indices suggested that a limited number of correlated residuals within the same latent dimensions could improve model fit. In this CFA section, item numbers refer to the final 17-item form; accordingly, residual covariances were added between i16 and i17, i7 and i8, and i4 and i5. These item pairs were conceptually similar and belonged to the same subdimensions. Therefore, the items were retained to preserve content coverage, whereas the residual covariances were reported transparently. Future validation studies should re-examine these item pairs and consider wording revisions if similar residual associations are replicated. The modified three-factor model showed improved fit, $\chi^2(113) = 155.300$, $p = .005$, CFI = .962, TLI = .955, IFI = .963, RMSEA = .031, 90% CI [.017, .042], and SRMR = .038. In this final model, standardized factor loadings ranged from .45 to .84 for Planning, .51 to .83 for Implementation, and .54 to .85 for Evaluation. Latent factor correlations ranged from .68 to .84, indicating that the three dimensions were related yet distinguishable.

For comparison purposes, a single-factor model was also tested. The unmodified single-factor model yielded weaker fit than the three-factor model, $\chi^2(119) = 314.200$, $p < .001$, CFI = .826, TLI = .801, IFI = .829, RMSEA = .065, 90% CI [.056, .073], and SRMR = .067. Although a modified single-factor model showed some improvement, $\chi^2(116) = 231.300$, $p < .001$, CFI = .897, TLI = .879, IFI = .899, RMSEA = .050, 90% CI [.041, .060], and SRMR = .056, it still fit the data more poorly than the modified three-factor solution. Taken together, these findings supported the three-factor structure of the scale.

Table 3
Comparison of Alternative Confirmatory Factor Models

Model	χ^2	df	χ^2/df	p	CFI	TLI	IFI	RMSEA	90% CI for RMSEA	SRMR
Three-factor, unmodified	218.700	116	1.886	< .001	.908	.893	.910	.047	[.038, .057]	.048
Three-factor, modified	155.300	113	1.374	.005	.962	.955	.963	.031	[.017, .042]	.038
Second-order, unmodified	218.700	116	1.886	< .001	.908	.893	.910	.047	[.038, .057]	.048
Second-order, modified	155.300	113	1.374	.005	.962	.955	.963	.031	[.017, .042]	.038
Single-factor, unmodified	314.200	119	2.641	< .001	.826	.801	.829	.065	[.056, .073]	.067
Single-factor, modified	231.300	116	1.994	< .001	.897	.879	.899	.050	[.041, .060]	.056

Note. The second-order models yielded the same global fit as their corresponding correlated three-factor models, indicating empirical equivalence at the level of global fit.

As shown in Table 3, the modified correlated three-factor and modified second-order models yielded equivalent global fit and both fit the data better than the single-factor alternatives. Although the modified single-factor model improved upon the unmodified single-factor solution, it still fit the data less well than the modified correlated three-factor and modified second-order solutions. These findings support the interpretation of self-regulated teaching as a multidimensional construct consisting of planning, implementation, and evaluation.

In addition to the correlated three-factor solution, a second-order CFA model was tested to examine whether the three first-order dimensions could be represented under a broader self-regulated teaching construct. The unmodified and modified second-order models yielded the same global fit as their corresponding correlated three-factor models. In the second-order solution, the standardized loadings of the first-order factors on the higher-order SRTS factor were .826 for Planning, .915 for Implementation, and .921 for Evaluation. Because the correlated three-factor and second-order models were empirically equivalent in terms of global fit, model choice was guided by interpretability rather than fit alone (Rodriguez, Reise and Haviland, 2016b). The correlated three-factor solution was retained as the primary representation because it was more consistent with the conceptualization of self-regulated teaching as three related but distinct phases. The bifactor-based analyses also indicated that the total score retained meaningful general-factor saturation ($\omega = .912$ in the CFA sample; $\omega h = .877$, ECV = .816, PUC = .640). Accordingly, the SRTS was interpreted primarily through its

Planning, Implementation, and Evaluation subdimensions, while the total score was treated as a complementary global indicator.

Additional internal-structure evidence was examined through convergent and discriminant validity analyses. Composite reliability coefficients were .780 for Planning, .928 for Implementation, and .891 for Evaluation. The corresponding AVE values were .552, .592, and .626, indicating acceptable convergent validity for all three dimensions. Discriminant validity was evaluated using HTMT. The HTMT values were .691 between Planning and Implementation, .668 between Planning and Evaluation, and .789 between Implementation and Evaluation. These values remained below commonly used critical thresholds, suggesting that the three dimensions were empirically distinguishable despite their substantial interrelatedness. At the same time, the relatively stronger association between Implementation and Evaluation indicated that these two dimensions are closely linked within the broader process of self-regulated teaching. Taken together, these results provide additional evidence for the internal structure of the SRTS. They indicate that the three dimensions are sufficiently related to represent a broader self-regulated teaching construct. At the same time, the findings support cautious interpretation of both the subdimension scores and the total score.

Table 4
Additional Evidence for the Internal Structure of the SRTS

Section	Variable / Comparison	N	ω	ωh	ECV	PUC	CR	AVE	r	HTMT
General factor	Total score, CFA sample	394	.912	.877	.816	.640				
	Total score, criterion sample	207	.917							
Convergent validity	Planning	394					.780	.552		
	Implementation	394					.928	.592		
	Evaluation	394					.891	.626		
Discriminant validity	Planning-Implementation	394							.757	.691
	Planning-Evaluation	394							.761	.668
	Implementation-Evaluation	394							.843	.789

Note. ω = McDonald's omega total; ωh = omega hierarchical; ECV = explained common variance; PUC = percentage of uncontaminated correlations; CR = composite reliability; AVE = average variance extracted; r = latent factor correlation; HTMT = heterotrait-monotrait ratio. Omega total values for the CFA and criterion-related validity samples were obtained in SPSS, and bifactor-based indices were estimated for the CFA sample.

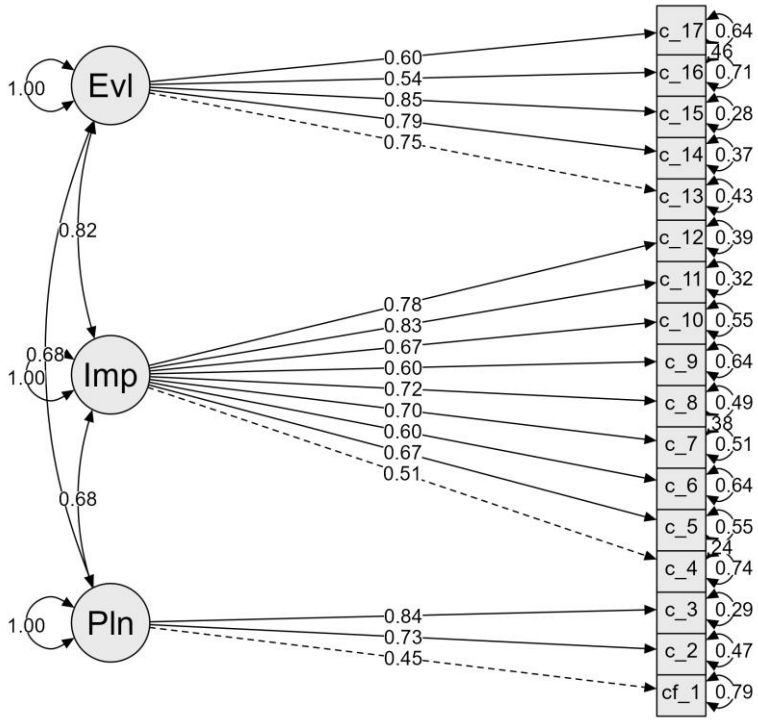


Figure 2. Final Standardized Three-Factor CFA Model

Measurement Invariance Across Gender

Measurement invariance across gender was examined using multigroup confirmatory factor analysis on the CFA sample. Configural, metric, and scalar invariance models were tested for the final three-factor structure. The configural model showed acceptable fit, $\chi^2(226) = 392.30$, CFI = .932, TLI = .918, RMSEA = .061, and SRMR = .052. Constraining factor loadings to equality across groups yielded the metric model, which also showed acceptable fit, $\chi^2(240) = 416.20$, CFI = .928, TLI = .918, RMSEA = .061, and SRMR = .069. Constraining both factor loadings and item intercepts yielded the scalar model, $\chi^2(257) = 431.80$, CFI = .928, TLI = .924, RMSEA = .059, and SRMR = .067. Changes in fit indices across nested models were small (configural to metric: ΔCFI = .004, ΔRMSEA = .000; metric to scalar: ΔCFI = .000, ΔRMSEA = .002), supporting configural, metric, and scalar invariance across gender. These findings suggest that the scale operates similarly for female and male teachers and that latent mean comparisons across gender are meaningful.

Criterion-Related Validity

After establishing additional evidence for the internal structure of the SRTS, criterion-related validity was examined using the Classroom Management Skills Scale (CMSS). As shown in Table 5, the Planning, Implementation, and Evaluation subdimensions were positively associated with the CMSS total score and its subdimension scores. In addition, the total SRTS score was strongly and positively correlated with the total CMSS score ($r = .751$, $p < .001$). Because the SRTS is primarily multidimensional, the subdimension-level associations were treated as the primary criterion-related validity evidence, whereas the total-score correlation was interpreted as a complementary indicator of the broader association between self-regulated teaching and classroom management.

Table 5
Correlations Between SRTS and CMSS along With Their Sub-Dimensions

	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**
1. PLN	1										
2. IMP	.646	1									
3. EVL	.537	.690	1								
4. SRTS	.771	.948	.851	1							
5. PHC	.617	.617	.592	.689	1						
6. EUT	.517	.597	.557	.644	.683	1					
7. PRS	.495	.574	.551	.624	.688	.627	1				
8. ATC	.459	.588	.563	.628	.613	.607	.681	1			
9. ASQ	.506	.646	.570	.674	.665	.602	.688	.720	1		
10. PSF	.438	.600	.547	.625	.636	.642	.821	.642	.716	1	
11. CMSS	.584	.700	.654	.751	.820	.809	.910	.817	.848	.888	1
M	12.87	39.04	21.55	73.46	9.98	13.09	19.42	10.08	13.48	12.63	78.68
SD	1.57	4.14	2.43	7.21	1.33	1.84	2.80	1.41	1.64	1.92	9.38
Skewness	-.23	-.38	-.12	-.26	.17	.07	.09	-.03	.15	.15	.28
Kurtosis	-.68	-.18	-.57	-.24	-.95	-.88	-1.02	-1.03	-1.01	-.74	-1.05

Note. PLN = Planning; IMP = Implementation; EVL = Evaluation; SRTS = Self-Regulated Teaching Scale total score; PHC = Physical Conditions; EUT = Efficient Use of Time; PRS = Presentation; ATC = Attendance in Class; ASQ = Asking Questions; PSF = Psychological Factors; CMSS = Classroom Management Skills Scale total score. All correlations were significant at $p < .001$.

Before the regression analysis, the relevant assumptions were examined in the criterion-related validity sample. Multicollinearity was not a concern, as tolerance values ranged from .42 to .57 and VIF values ranged from 1.77 to 2.40. The independence of errors assumption was supported by the Durbin-Watson value of 1.776. The standardized residuals were approximately normally distributed (skewness = -.010, kurtosis = -.342; Kolmogorov-Smirnov = .050, p = .200; Shapiro-Wilk = .994, p = .593), and their values ranged from -2.71 to 2.47. Scatterplots did not indicate serious deviations from linearity or homoscedasticity. To complement the correlation findings, a multiple regression analysis was conducted to examine the unique contributions of the SRTS subdimensions to CMSS scores.

Table 6
Multiple Regression Results Predicting CMSS from SRTS Subdimensions

Predictors	**B**	**SE B**	**Beta**	**t**	**p**	**Tolerance**	**VIF**
Planning	1.07	.37	.18	2.92	.004	.57	1.77
Implementation	.86	.16	.38	5.31	< .001	.42	2.40
Evaluation	1.14	.25	.30	4.55	< .001	.51	1.96

Note. Criterion variable = CMSS total score. $R^2 = .565$, $F(3, 203) = 87.77$, $p < .001$. B = unstandardized coefficient; SE B = standard error; Beta = standardized coefficient; VIF = variance inflation factor.

A multiple regression analysis further showed that Planning ($\beta = .18$, $p = .004$), Implementation ($\beta = .38$, $p < .001$), and Evaluation ($\beta = .30$, $p < .001$) significantly predicted classroom management skills. The model was statistically significant, $F_{(3, 203)} = 87.77$, $p < .001$, and explained 56.5% of the variance in CMSS scores ($R^2 = .565$). In addition, tolerance values ranged from .42 to .57 and VIF values ranged from 1.77 to 2.40, indicating no serious multicollinearity problem. Taken together, these findings supported the criterion-related validity of the scale.

Reliability

Reliability was examined using internal consistency, model-based reliability, and test-retest stability. Because the SRTS is multidimensional, reliability was interpreted primarily at the subdimension level, whereas total-score coefficients were treated as supplementary and evaluated together with model-based indices.

In the CFA sample (n = 394), Cronbach's alpha coefficients were .689 for Planning, .885 for Implementation, and .843 for Evaluation. In the criterion-related validity sample (n = 207), the corresponding alpha coefficients were .623, .885, and .838. These results indicate acceptable to high

internal consistency, although the relatively lower alpha for Planning should be interpreted in light of the small number of items in that factor.

For the total score, McDonald's omega was .912 in the CFA sample and .917 in the criterion-related validity sample. Bifactor-based indices also supported meaningful general-factor saturation (ωh = .877, ECV = .816, PUC = .640). Test-retest reliability, assessed in a subsample of 93 teachers over a 4-6-week interval, yielded ICC values of .780 for Planning, .820 for Implementation, and .800 for Evaluation. Corrected item-total correlations were generally acceptable, although Item 1 showed a relatively lower value and was retained because it represented an essential aspect of the planning phase. These results indicate that the SRTS is best interpreted primarily at the subdimension level, while the total score may be used cautiously as a complementary global indicator.

Table 7
Reliability Evidence for the SRTS

Scale	k	α394	α207	ω394	ω207	CITC394	CITC207	ICC (95% CI)
Planning	3	.689	.623			.35-.63	.33-.67	.780 [.660, .850]
Implementation	9	.885	.885			.48-.76	.51-.74	.820 [.720, .880]
Evaluation	5	.843	.838			.50-.74	.47-.69	.800 [.700, .870]
Total	17	.914	.918	.912	.917	.35-.76	.33-.74	

Note. α394 = Cronbach's alpha in the CFA sample (n = 394); α207 = Cronbach's alpha in the criterion-related validity sample (n = 207); ω394 = McDonald's omega in the CFA sample; ω207 = McDonald's omega in the criterion-related validity sample; CITC394 = corrected item-total correlation range in the CFA sample; CITC207 = corrected item-total correlation range in the criterion-related validity sample; ICC = test-retest intraclass correlation coefficient based on a subsample of 93 teachers. Because the SRTS is multidimensional, total-score coefficients should be interpreted together with model-based indices rather than alpha alone. Values in the Total row summarize the overall scale, whereas ICC values are reported for subdimensions.

Discussion, Conclusion, and Recommendations

The present study aimed to develop a theoretically grounded and psychometrically supported instrument for assessing teachers' self-regulated teaching skills on the basis of Zimmerman's cyclical self-regulated learning model. Across the qualitative and quantitative phases, a consistent three-dimensional pattern emerged, consisting of Planning, Implementation, and Evaluation. This structure is theoretically meaningful because it corresponds to teachers' pre-instructional, during-instructional, and post-instructional regulation of teaching practice.

The CFA findings showed that the modified correlated three-factor and modified second-order models fit the data better than the single-factor alternatives and yielded equivalent global fit. Therefore, model selection was guided by interpretability rather than fit alone. The correlated three-factor representation was retained as primary because it was more consistent with the conceptualization of self-regulated teaching as three related but distinguishable phases. At the same time, model-based indices indicated that the total score retained meaningful psychometric value. Thus, the SRTS should be interpreted primarily through its subdimension scores, whereas the total score may be used cautiously as a complementary summary indicator when a broader interpretation is theoretically justified. Additional internal-structure analyses were consistent with this interpretation. The AVE and CR values indicated acceptable convergent evidence, and the HTMT values supported discriminant validity at an acceptable level. The relatively stronger association between Implementation and Evaluation was theoretically plausible because teachers' instructional enactment and reflective appraisal may overlap in classroom practice. Accordingly, the dimensions should be viewed not as independent domains, but as related and distinguishable components of a broader self-regulatory teaching process.

The SRTS was also positively associated with classroom management skills. Correlation and regression findings showed that Planning, Implementation, and Evaluation each contributed to classroom management. This result supports the criterion-related validity of the scale and suggests that self-regulated teaching may function as a professional resource for effective classroom practice.

Because the scale is primarily multidimensional, the subdimension-level findings provide the main validity evidence, while the total-score association with CMSS should be interpreted as complementary.

Several limitations should be noted. First, both the SRTS and CMSS were administered as self-report measures, so common method bias cannot be completely ruled out. Second, the final CFA model included a limited number of theoretically defensible correlated residuals within dimensions. Third, the Planning subdimension showed relatively lower alpha values and should be re-examined in future validation studies. Fourth, the empirical equivalence of the correlated three-factor and second-order models indicates that the structure should be replicated in independent samples. Fifth, although the total score was supported by model-based evidence, the scale remains primarily multidimensional. Finally, measurement invariance was tested only across gender; therefore, future studies should examine invariance across school level, teaching experience, and subject area.

In conclusion, the findings indicate that the Self-Regulated Teaching Scale is a theoretically grounded and psychometrically supported instrument for assessing teachers' self-regulated teaching. The scale is best used primarily through its Planning, Implementation, and Evaluation subdimension scores, with the total score serving as a complementary global indicator when needed. Future research should replicate the findings in independent teacher samples, compare SRTS scores with external indicators such as classroom observation, peer evaluation, administrator ratings, or student outcomes, and re-examine the item pairs that required correlated residuals. Further studies should also continue to evaluate the dimensionality of the total score using contemporary psychometric approaches.

Contribution of Researchers

Since the study is single-authored, the contribution rate is 100%.

Conflict of Interest

The author has no conflict of interest.

References

Aslanderen, M., Çilsalar Sagnak, H. ve Sağlam, M. (2024). Okul öncesi öğretmenliği programındaki öğretim teknolojileri dersinin yansıtıcı değerlendirme modeli ile değerlendirilmesi. *Ahi Evran Üniversitesi Sosyal Bilimler Enstitüsü Dergisi, 10*(3), 543-562. https://doi.org/10.31592/aeusbed.1433721

Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory*. Englewood Cliffs, NJ: Prentice Hall.

Banks, R. S. and Thompson, C. L. (1995). *Educational psychology: For teachers in training*. Eagan: West Publishing Company.

Beavers, A. S., Lounsbury, J. W., Richards, J. K., Huck, S. W., Skolits, G. J. and Esquivel, S. L. (2013). Practical considerations for using exploratory factor analysis in educational research. *Practical Assessment, Research, and Evaluation, 18*(1), 6. https://doi.org/10.7275/qv2q-rk76

Boekaerts, M. (1997). Self-regulated learning: A new concept embraced by researchers, policy makers, educators, teachers, and students. *Learning and Instruction, 7*(2), 161-186. https://doi.org/10.1016/S0959-4752(96)00015-1

Borkowski, J. G. (1996). Metacognition: Theory or chapter heading? *Learning and Individual Differences, 8*(4), 391-402. https://doi.org/10.1016/S1041-6080(96)90025-4

Brown, T. A. (2015). *Confirmatory factor analysis for applied research* (2nd ed.). New York: The Guilford Press.

Büyükkaragöz, S. S. ve Çivi, C. (1999). *Genel öğretim metotları* (10. baskı). İstanbul: Beta Basım Yayın.

Carpenter, S. (2018). Ten steps in scale development and reporting: A guide for researchers. *Communication Methods and Measures, 12*(1), 25-44. https://doi.org/10.1080/19312458.2017.1396583

Christie, F. (2014). Curriculum genres: Planning for effective teaching. In N. Hall, J. Larson and J. Marsh (Eds.), *The powers of literacy* (RLE Edu I, pp. 154-178). New York: Routledge.

Cothran, D. J. and Kulinna, P. H. (2008). Teachers' knowledge about and use of teaching models. *Physical Educator, 65*(3), 122-133.

Creswell, J. W. (2017). *Research design: Qualitative, quantitative, and mixed methods approaches* (4th ed.). London: Sage Publications.

Davis, L. L. (1992). Instrument review: Getting the most from a panel of experts. *Applied Nursing Research, 5*(4), 194-197. https://doi.org/10.1016/S0897-1897(05)80008-4

Delson-Karan, M. (1984). A three-dimensional approach to evaluating teaching performance. *Canadian Modern Language Review, 40*(3), 414-428. https://doi.org/10.3138/cmlr.40.3.414

Demirel, Ö. (2009). *Planlamadan değerlendirmeye öğretme sanatı* (19. baskı). Ankara: Pegem Akademi.

DeVellis, R. F. (2017). *Scale development: Theory and application* (4th ed.). Thousand Oaks, CA: Sage Publications.

Donche, V., De Maeyer, S., Coertjens, L., Van Daal, T. and Van Petegem, P. (2013). Differential use of learning strategies in first-year higher education: The impact of personality, academic motivation, and teaching strategies. *British Journal of Educational Psychology, 83*(2), 238-251. https://doi.org/10.1111/bjep.12016

El-Adl, A. and Alkharusi, H. (2020). Relationships between self-regulated learning strategies, learning motivation and mathematics achievement. *Cypriot Journal of Educational Sciences, 15*(1), 104-111. https://doi.org/10.18844/cjes.v15i1.4461

Ergün, M., Ergezer, B., Çevik, İ. ve Özdaş, A. (1999). *Öğretmenlik mesleğine giriş*. Ankara: Ocak Yayınları.

Filice, S., Tregunno, D., Edge, D. and Egan, R. (2020). Re-imaging clinical education: The interdependence of the self-regulated clinical teacher and nursing student. *International Journal of Nursing Education Scholarship, 17*(1), 20190056. https://doi.org/10.1515/ijnes-2019-0056

Garson, G. D. (2013). *Validity & reliability* (2013 ed.). Asheboro: Statistical Associates Publishers.

Grant, L., Hindman, J. and Stronge, J. (2013). *Planning, instruction, and assessment: Effective teaching practices*. New York: Routledge.

Güneş, G. ve Demir, M. K. (2025). Sınıf öğretmenlerinin öğretim programlarını eleştirel okuma açısından yeterli bulma durumları. *Ahi Evran Üniversitesi Sosyal Bilimler Enstitüsü Dergisi, 11*(2), 711-727. https://doi.org/10.31592/aeusbed.1692182

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

Karakelle, S. (2005). Analyzing teachers' definitions of effective teachers according to effective teaching dimensions. *Education and Science, 30*(135), 1-10.

Kılıç, A. (2006). Levels of observation and practice of teaching skills of student-teachers. *MANAS Journal of Social Studies, 8*(16), 155-168.

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

Kurt, Ş. H. ve Sığırtmaç, A. D. (2021). The effect of practices supporting self-regulated learning and preschool teachers' self-regulations on classroom management skills. *Journal of Early Childhood Studies, 5*(1), 135-151. https://doi.org/10.24130/eccd-jecs.1967202151245

Lee, M., Lee, S. Y., Kim, J. E. and Lee, H. J. (2023). Domain-specific self-regulated learning interventions for elementary school students. *Learning and Instruction, 88*, 101810. https://doi.org/10.1016/j.learninstruc.2023.101810

Lesh, R. and Lehrer, R. (2003). Models and modeling perspectives on the development of students and teachers. *Mathematical Thinking and Learning, 5*(2-3), 109-129. https://doi.org/10.1080/10986065.2003.9679996

Liu, S. and Meng, L. (2009). Perceptions of teachers, students and parents of the characteristics of good teachers: A cross-cultural comparison of China and the United States. *Educational Assessment, Evaluation and Accountability, 21*, 313-328. https://doi.org/10.1007/s11092-009-9077-z

MacCallum, R. C., Widaman, K. F., Zhang, S. and Hong, S. (1999). Sample size in factor analysis. *Psychological Methods, 4*(1), 84-99. https://doi.org/10.1037/1082-989X.4.1.84

Miles, M. B., Huberman, A. M. and Saldaña, J. (2014). *Qualitative data analysis: An expanded sourcebook* (3rd ed.). London: Sage Publications.

Peeters, J., De Backer, F., Reina, V. R., Kindekens, A., Buffel, T. and Lombaerts, K. (2014). The role of teachers' self-regulatory capacities in the implementation of self-regulated learning practices. *Procedia - Social and Behavioral Sciences, 116*, 1963-1970. https://doi.org/10.1016/j.sbspro.2014.01.504

Pintrich, P. R. (2000). The role of goal orientation in self-regulated learning. In M. Boekaerts, P. R. Pintrich and M. Zeidner (Eds.), *Handbook of self-regulation* (pp. 451-502). San Diego, CA: Academic Press. https://doi.org/10.1016/B978-012109890-2/50043-3

Randi, J. (2004). Teachers as self-regulated learners. *Teachers College Record, 106*(9), 1825-1853. https://doi.org/10.1111/j.1467-9620.2004.00407.x

Rattray, J. and Jones, M. C. (2007). Essential elements of questionnaire design and development. *Journal of Clinical Nursing, 16*(2), 234-243. https://doi.org/10.1111/j.1365-2702.2006.01573.x

Rodriguez, A., Reise, S. P. and Haviland, M. G. (2016a). Applying bifactor statistical indices in the evaluation of psychological measures. *Journal of Personality Assessment, 98*(3), 223-237. https://doi.org/10.1080/00223891.2015.1089249

Rodriguez, A., Reise, S. P. and Haviland, M. G. (2016b). Evaluating bifactor models: Calculating and interpreting statistical indices. *Psychological Methods, 21*(2), 137-150. https://doi.org/10.1037/met0000045

Seçer, İ. (2018). *Psikolojik test geliştirme ve uyarlama süreci: SPSS ve LISREL uygulamaları* (2. baskı). Ankara: Anı.

Singh, K., Junnarkar, M. and Kaur, J. (2016). *Measures of positive psychology: Development and validation*. Berlin: Springer.

Sönmez, V. (2020). *Program geliştirmede öğretmen el kitabı* (19. baskı). Ankara: Anı.

Strong, M., Gargani, J. and Hacifazlioğlu, Ö. (2011). Do we know a successful teacher when we see one? Experiments in the identification of effective teachers. *Journal of Teacher Education, 62*(4), 367-382. https://doi.org/10.1177/0022487110390221

Stronge, J. H. (2018). *Qualities of effective teachers* (3rd ed.). Alexandria, VA: ASCD.

Sutton, R. E., Mudrey-Camino, R. and Knight, C. C. (2009). Teachers' emotion regulation and classroom management. *Theory into Practice, 48*(2), 130-137. https://doi.org/10.1080/00405840902776418

Tuckman, B. W. (1995). Assessing effective teaching. *Peabody Journal of Education, 70*(2), 127-138. https://doi.org/10.1080/01619569509538827

Watkins, M. W. (2021). *A step-by-step guide to exploratory factor analysis with R and RStudio*. New York: Routledge.

Worthington, R. L. and Whittaker, T. A. (2006). Scale development research: A content analysis and recommendations for best practices. *The Counseling Psychologist, 34*(6), 806-838. https://doi.org/10.1177/0011000006288127

Yalçınkaya, M. ve Tonbul, Y. (2002). The perception and observation regarding to the primary school teachers' classroom management skills. *Ege Journal of Education, 1*(2), 1-10.

Zheng, B., Ward, A. and Stanulis, R. (2020). Self-regulated learning in a competency-based and flipped learning environment: Learning strategies across achievement levels and years. *Medical Education Online, 25*(1), 1686949. https://doi.org/10.1080/10872981.2019.1686949

Zimmerman, B. J. (2000). Attaining self-regulation: A social cognitive perspective. In M. Boekaerts, P. R. Pintrich and M. Zeidner (Eds.), *Handbook of self-regulation* (pp. 13-39). San Diego, CA: Academic Press. https://doi.org/10.1016/B978-012109890-2/50031-7

Zimmerman, B. J. (2013). From cognitive modeling to self-regulation: A social cognitive career path. *Educational Psychologist, 48*(3), 135-147. https://doi.org/10.1080/00461520.2013.794676

Zimmerman, B. J. and Campillo, M. (2003). Motivating self-regulated problem solvers. In J. E. Davidson and R. J. Sternberg (Eds.), *The psychology of problem solving* (pp. 233-262). Cambridge: Cambridge University Press.

Genişletilmiş Özet

Bu çalışma, öğretmenlerin öğretim sürecini planlama, uygulama ve değerlendirme aşamalarında ne ölçüde öz düzenlemeli biçimde yönettiklerini belirlemeye yönelik bir ölçme aracı geliştirmeyi amaçlayan kuramsal ve psikometrik bir araştırmadır. Öz düzenlemeli öğretim, öğretmenin yalnızca ders içi performansını değil, öğretim öncesindeki hazırlıklarını, öğretim sırasında hedeflere uygun biçimde hareket etmesini ve öğretim sonrasında kendi uygulamalarını eleştirel olarak değerlendirmesini kapsayan döngüsel bir süreç olarak ele alınmıştır. Bu yönüyle araştırma, Zimmerman'ın öz düzenlemeli öğrenme modelini öğretmenlerin mesleki öğretim davranışlarına uyarlayarak alanyazında öğretmen yeterlikleri, sınıf yönetimi ve genel öz düzenleme ölçeklerinden farklılaşan özgül bir ölçme aracı ortaya koymayı hedeflemektedir. Araştırmada önce odak grup görüşmeleriyle öğretmenlerin öğretim öncesi, öğretim sırası ve öğretim sonrası davranışlarına ilişkin nitel veriler toplanmış, bu verilerden hareketle madde havuzu oluşturulmuş ve maddeler uzman görüşü ile öğretmen değerlendirmelerine göre gözden geçirilmiştir. Ardından farklı öğretmen örneklemlerinden elde edilen verilerle açımlayıcı faktör analizi, doğrulayıcı faktör analizi, ölçüt ilişkili geçerlik, iç tutarlılık, model temelli güvenirlik ve test-tekrar test analizleri yürütülmüştür. Elde edilen bulgular, ölçeğin Planlama, Uygulama ve Değerlendirme olmak üzere üç boyutlu ve 17 maddelik bir yapı sergilediğini, bu yapının kuramsal modelle uyumlu olduğunu ve sınıf yönetimi becerileriyle anlamlı ilişkiler gösterdiğini ortaya koymuştur. Bulgular aynı zamanda ölçeğin alt boyut puanları üzerinden yorumlanmasının daha uygun olduğunu, toplam puanın ise ancak genel bir göstergeye ihtiyaç duyulduğunda tamamlayıcı biçimde kullanılabileceğini göstermektedir. Bu nedenle genişletilmiş özet, ölçeğin geliştirilme gerekçesini, geçerlik ve güvenirlik kanıtlarının kapsamını ve kullanım biçimini okuyucuya açık biçimde aktarmayı amaçlamaktadır. Böylece okuyucu, ana metindeki ayrıntılı analizlere geçmeden önce araştırmanın problemi, yöntemi, temel sonuçları ve önerileri hakkında bütünlüklü bir çerçeve edinebilecektir. Bu genişletilmiş özet, çalışmanın gerekçesini, yöntemsel aşamalarını, temel bulgularını, ölçeğin nasıl yorumlanması gerektiğini ve gelecekte yapılabilecek araştırmalara yönelik önerileri ana metindeki akışı koruyarak sunmaktadır.

Giriş

Eğitim-öğretim hizmetleri, öğrenci, öğretmen ve programın etkileşimine dayalı çok boyutlu bir yapıdır. Bu yapı içinde öğretmen, yalnızca bilgiyi aktaran kişi değil; öğretimi planlayan, uygulayan, değerlendiren ve gerektiğinde yeniden düzenleyen temel aktördür. Bu nedenle onların öğretim sürecini nasıl yönettikleri, öğrencilerin öğrenme yaşantıları ve akademik başarıları açısından önem taşımaktadır. Öz düzenleme, bireyin belirli amaçlara ulaşmak için düşünce, duygu ve davranışlarını bilinçli biçimde izlemesi, yönlendirmesi ve değerlendirmesi olarak tanımlanabilir. Zimmerman'ın döngüsel öz düzenlemeli öğrenme modeli de bu süreci ön düşünme, performans ve öz yansıtma aşamalarıyla açıklamaktadır. Bu çalışmada öğretmenlerin öğretim faaliyetlerindeki öz düzenleyici becerileri, Zimmerman'ın modeline dayalı planlama, uygulama ve değerlendirme boyutlarıyla ele alınmıştır. Alanyazında öğretmen yeterlikleri, sınıf yönetimi ve genel öz düzenleme ile ilgili ölçme araçları bulunmakla birlikte, öğretmenlerin öz düzenlemeli öğretim uygulamalarını doğrudan bu kuramsal çerçeveye dayalı olarak ölçen kısa ve geçerli bir aracın sınırlı olduğu gerekçesinden ve ihtiyacından hareketle, mevcut araştırmanın amacı, öğretmenlerin öz düzenlemeli öğretim becerilerini ölçmeye yönelik geçerli ve güvenilir bir ölçek geliştirmektir.

Yöntem

Araştırmada keşfedici sıralı karma yöntem deseni kullanılmıştır. İlk aşamada, 12 öğretmenle gerçekleştirilen üç odak grup görüşmesinden elde edilen nitel veriler doğrultusunda madde havuzu oluşturulmuştur. Görüşmeler öğretmenlerin öğretim öncesinde, öğretim sırasında ve öğretim sonrasında yaptıkları uygulamalara odaklanmış; elde edilen veriler üç bağımsız kodlayıcı tarafından incelenmiş ve kodlar planlama, uygulama ve değerlendirme temaları altında düzenlenmiş ve Zimmerman'ın modeliyle uyumlu bir kavramsal yapı oluşturulmuştur. Başlangıçta hazırlanan 32 maddelik havuz, dil ve alan uzmanlarının görüşleri doğrultusunda gözden geçirilmiş; kapsam geçerlik

katsayısı düşük olan iki madde çıkarılmıştır. Daha sonra 30 maddelik taslak form, 75 öğretmen tarafından açıklık ve anlaşılırlık açısından değerlendirilmiştir.

Nicel analizler beş bağımsız örneklem üzerinde yürütülmüştür. Açımlayıcı faktör analizi 226 öğretmenden elde edilen verilerle, doğrulayıcı faktör analizi ve iç tutarlılık analizleri 394 öğretmenden elde edilen verilerle, ölçüt ilişkili geçerlik analizleri ise 207 öğretmenden elde edilen verilerle gerçekleştirilmiştir. Ayrıca 93 öğretmenden oluşan bir alt örneklem üzerinde 4-6 hafta arayla test-tekrar test analizi yapılmıştır. AFA'da temel eksen faktörleştirme ve Direct Oblimin döndürme yöntemi kullanılmıştır. DFA, JASP ile yapılmış, model uyumu χ^2/sd, CFI, TLI, IFI, RMSEA ve SRMR ile değerlendirilmiştir. Ölçüt ilişkili geçerlik için Sınıf Yönetimi Becerileri Ölçeği kullanılmıştır. İç tutarlılık, Cronbach alfa, McDonald omega, model temelli güvenirlik göstergeleri ve test-tekrar test katsayılarıyla birlikte incelenmiştir.

Bulgular

AFA'da Planlama, Uygulama ve Değerlendirme olarak adlandırılan üç faktörlü ve 17 maddelik bir yapı elde edilmiştir. Faktör yapısı, Zimmerman'ın döngüsel öz düzenleme modelinin öğretim sürecine uyarlanmış biçimiyle kuramsal olarak uyumludur. DFA, modifikasyonlu üç faktörlü modelin veriye iyi uyum sağladığını (χ^2/sd = 1.374, CFI = .962, TLI = .955, IFI = .963, RMSEA = .031, SRMR = .038), tek faktörlü modellerin üç faktörlü modele göre daha zayıf uyumlu olduğunu göstermiştir. Korelasyonel üç faktörlü model ile ikinci düzey modelin küresel uyum değerleri eşdeğer bulunmuş; bu nedenle model seçimi yalnızca uyum indekslerine değil, kuramsal yorumlanabilirliğe dayandırılmıştır. Ölçek, temel olarak Planlama, Uygulama ve Değerlendirme alt boyutları üzerinden yorumlanmış; toplam puan ise tamamlayıcı bir genel gösterge olarak ele alınmıştır.

Ek model temelli bulgular da toplam puanın sınırlı fakat anlamlı biçimde kullanılabileceğini göstermiş, McDonald omega, omega hiyerarşik, açıklanan ortak varyans ve kirlenmemiş korelasyon yüzdesi değerleri toplam puanın genel öz düzenlemeli öğretim göstergesi olarak dikkatli biçimde yorumlanabileceğini desteklemiştir. Ayrıca AVE, CR ve HTMT değerleri yakınsak ve ayırt edici geçerlik açısından kabul edilebilir sonuçlar sunmuştur. Ölçüt ilişkili geçerlik kapsamında, Öz Düzenlemeli Öğretim Becerileri Ölçeği'nin boyutları ve toplam puanı ile Sınıf Yönetimi Becerileri Ölçeği puanları arasında pozitif ve anlamlı ilişkiler bulunmuş, Planlama, Uygulama ve Değerlendirmenin sınıf yönetimi becerilerini anlamlı biçimde yordadığı görülmüştür. Güvenirlik analizlerinde boyutlara ait iç tutarlılık katsayıları kabul edilebilir ila yüksek düzey arasında değişmiş, toplam puan için omega katsayıları yüksek bulunmuş ve test-tekrar test sonuçları ölçeğin zamansal kararlılığını desteklemiştir.

Sonuç, Tartışma ve Öneriler

Araştırma bulguları, Öz Düzenlemeli Öğretim Becerileri Ölçeği'nin öğretmenlerin öğretim sürecindeki öz düzenleyici becerilerini ölçmek için geçerli ve güvenilir bir araç olduğunu göstermektedir. Ölçek, öğretimin planlama, uygulama ve değerlendirme aşamalarını kapsayan üç boyutlu bir yapı sunmaktadır. Bu yapı, öğretmenlerin öğretim sürecini yalnızca uygulama düzeyinde değil, ön hazırlık ve son değerlendirme boyutlarıyla birlikte ele almaya olanak sağlamaktadır. Ölçeğin alt boyut puanlarının öncelikli olarak kullanılması, toplam puanın ise genel bir göstergeye ihtiyaç duyulduğunda tamamlayıcı biçimde yorumlanması önerilmektedir. Bununla birlikte, çalışmanın bazı sınırlılıkları vardır. Veriler öz bildirim yoluyla toplanmıştır, bazı madde çiftlerinde ilişkili hata kovaryansları modele eklenmiştir ve ölçme değişmezliği yalnızca cinsiyet açısından incelenmiştir. Gelecek araştırmalarda ölçeğin farklı örneklemlerde, farklı okul düzeylerinde, branşlarda ve öğretmen deneyimi gruplarında yeniden test edilmesi; sınıf gözlemi, akran değerlendirmesi, yönetici görüşleri ya da öğrenci çıktıları gibi dış ölçütlerle karşılaştırılması önerilmektedir. Ayrıca ilişkili hata kovaryansı gerektiren madde çiftlerinin yeni çalışmalarda yeniden incelenmesi ve toplam puanın boyutsallığının çağdaş psikometrik yaklaşımlarla test edilmeye devam edilmesi yararlı olacaktır.