Journal of Pedagogical Research
Volume 8, Issue 4, 2024
https://doi.org/10.33902/JPR.202429249

Research Article

Development of scale to measure teachers' curriculum assessment

Fatmir Mehmeti[1], Artan Reshani[2] and Erdoğan Tezci[3]

[1]*University "Ukshin Hoti" Prizren, Faculty of Education, Republic of Kosovo (ORCID: 0000-0002-9094-6654)*
[2]*University "Ukshin Hoti" Prizren, Faculty of Education, Republic of Kosovo (ORCID: 0000-0001-8124-0003)*
[3]*University of Balıkesir, Faculty of Education, Türkiye (ORCID: 0000-0003-2055-0192)*

Curriculum guides education and ensures unity and integrity in educational and training activities. Furthermore, it plays a crucial role in providing quality education to students. Education quality is therefore influenced by the quality of the curriculum. In addition to the quality of the curriculum design, its applicability in the classroom is equally important. Since teachers implement the curriculum in class, they are one of the best judges of the curriculum's quality, its applicability, its design, and the benefits derived from implementing it. It is important to assess the quality of a curriculum based on teachers' evaluations. There are numerous studies on teacher program evaluation in the literature, but it is also clear that a comprehensive curriculum evaluation scale is needed. To meet this need, a scale was developed based on teachers' evaluations. Taking into account the quality and elements of the curriculum, a two-dimensional structure was developed. An independent sample of 279 teachers for exploratory factor analysis (EFA) and 220 teachers for confirmatory factor analysis (CFA) participated in the study. Based on the EFA, 31 items and 8 factor structures were identified. CFA results showed adequate fit indices for the 8-factor structure. As evidence of construct validity, the scale demonstrated convergent and discriminant validity. The Cronbach's alpha and omega reliability coefficients were sufficient for reliability, and the items were discriminatory. The scale was found to be valid and reliable enough to assess the quality of the curriculum based on teachers' views. The scale will contribute in one aspect to assessing the curriculum, and in another aspect to evaluating the outcomes of teacher preparation and in-service training.

Keywords: Curriculum; Curriculum assessment scale; Teacher education; Scale development

Article History: Submitted 1 April 2024; Revised 28 June 2024; Published online 25 August 2024

1. Introduction

Curriculum serve as the foundation for reform efforts aimed at achieving social change (Goodlad, 1964; Macdonald, 2003). Since 2011, Kosovo has implemented significant curriculum reforms to modernize its pre-university education system. These reforms aim to move from goal-oriented, content-intensive education to a more dynamic, competency-based approach. This approach focuses on developing key skills and abilities students need to succeed in the 21st century, including critical thinking, problem solving, collaboration and creativity. The curriculum emphasizes clear learning outcomes expected of students at each grade level. These outcomes are intended to ensure that students acquire essential skills and knowledge relevant to their personal

Address of Corresponding Author

Artan Reshani, PhD, "Ukshin Hoti" University, Rruga e Shkronjave, nr. 1 Prizren 20000, Republic of Kosova.

✉ artan.reshani@uni-prizren.com

How to cite: Mehmeti, F., Reshani, A., & Tezci, E. (2024). Development of scale to measure teachers' curriculum assessment. *Journal of Pedagogical Research, 8*(4), 66-89. https://doi.org/10.33902/JPR.202429249

and professional lives. The curriculum promotes student-centered learning, where students are actively involved in their learning process. This includes more interactive and participatory teaching methods such as group work, projects and practical activities that promote deeper understanding and application of knowledge. Subjects are now taught in a more integrated manner, highlighting the connections between different areas of knowledge. This interdisciplinary approach helps students see the relevance of what they are learning and how it applies to real-world situations.

To ensure that these curriculum reforms effectively achieve their goals, curriculum assessment is a crucial aspect of educational development. Different researchers have developed different approaches to curriculum evaluation, recognizing the importance of involving partners in the process. Among these participants, teachers play a particularly important role because of their direct role in implementing the curriculum in the classroom. Their contributions to curriculum planning, implementation and evaluation are significant (Mehmeti & Tezci, 2008). Even the best-designed curriculum cannot achieve its intended outcomes if teachers do not fully understand it and apply it effectively. According to Dindar and Yaygın (2007), teachers must have a comprehensive understanding of the curriculum as well as the skills to implement and evaluate it appropriately. Therefore, teachers are considered primary sources in evaluating the performance and impact of a curriculum on students (Button, 2021). For that reason, several curriculum assessment models emphasize the importance of teacher involvement in curriculum evaluation. For example, the model developed by Metfessel and Michael consists of eight phases and involves teachers, administrators, students, and ordinary citizens either directly or indirectly in the assessment process. This comprehensive approach ensures that different perspectives are considered, resulting in a more holistic assessment (Stufflebeam, 2001). Similarly, Stake's participant-centered evaluation model focuses on gathering opinions from various curriculum partners, including teachers, students, parents, and administrators, and uses multiple data sources to evaluate the suitability and effectiveness of the curriculum (Kelly, 2004).

Given the central role of teachers in the process of curriculum implementation, their perspectives are crucial for thorough and accurate evaluation of curriculum. Teachers select and apply appropriate teaching methods in the classroom. Designs and plans instruction according to student needs. Assessing the students' progress. Teachers are those responsible for successfully implementing curriculum innovations. There are many studies in the literature addressing the role of teachers in curriculum implementation and development (Carless, 1998; Connelly, 1980; Fu, & Sibert, 2017; Wang & Cheng, 2009). Since teachers are the ones who implement the curriculum in the classroom, they are the ones who can best evaluate it. Marsh and Willis (2007) emphasized the difference between the planned/designed curriculum and the implemented curriculum and stated that different situations may arise during the implementation phase of the curriculum. It seems difficult for teachers who do not evaluate the curriculum as qualified to implement the planned/designed curriculum according to the understanding in the planned. In both cases, it can be said that the curriculum is difficult to implement effectively in the classroom (Burul & Tezci, 2022). Therefore, the success of the curriculum depends on how teachers evaluate it (Elliott, 1994; Evans, 1986; Karakuş, 2021). Although there are many studies on curriculum evaluation in the literature (Dagenais et al., 2003; Maren et al., 2021), it has been observed that there are not sufficient instrument for comprehensive curriculum assessment. The studies are generally qualitative, elements of curriculum or curriculum content or problems about implementations (Akıncı & Köse, 2021; Apsari, 2018; Ben- Chaim et al., 1994; Kern et al., 2007). The lack of a comprehensive measurement tool that covers the curriculum from the teacher perspective, from approach to content, from evaluations to recommended methods, is another issue that limits the studies. Carrying out the curriculum evaluation from the teachers' perspective will, on the one hand, provide information about the applicability of the curriculum in the classroom and, on the other hand, help obtain information about the teachers' competence. In this context, it is aimed to develop a sufficiently reliable and valid an instrument that includes teacher evaluation. It is

evaluated that this instrument will contribute to practitioners, researchers, curriculum developers and policy makers.

2. Literature Review

2.1. Curriculum Evaluation

Different researchers have developed various approaches for curriculum evaluation. It is a well-established fact that it is impossible to assert that any one of the evaluation approaches or models depicted is the best or most accurate. Each has its strengths and weaknesses and focuses on different elements in the evaluation process to obtain data. The choice of evaluation model and approach depends on what the evaluation aims to illuminate or why it is being conducted. Each approach offers distinct perspectives and methodologies, emphasizing different aspects of the evaluation process to obtain relevant data and insights. For examples, the goal of the goal-oriented evaluation approach is to ascertain the degree to which goals have been accomplished. Tayley, Metfessel-Michael, Provus, Bennett, Hammond, and Hammond are among its supporters (Brown, 1995). The primary goal of the systems-oriented assessment approach is to provide decision makers with the necessary information. Key figures in this approach include Stufflebeam, Dick and Carey, Kirkpatrick, Alkin, Saylor, Alexander and Lewis (Saylor & Alexander, 1973). The competency-based assessment approach focuses on professional practitioners evaluating a specific work. Its pioneer is Eisner (Şahan, 2007). The participant-oriented evaluation approach, led by Stake, Paris, and Hamilton, emphasizes the inclusion of participants and is crucial for determining evaluation criteria, needs, data, values, and outcomes (Ornstein & Hunkins, 2003). The adversary-oriented evaluation approach includes Wolf's evaluation model and aims to decide whether the model should continue based on different perspectives of evaluation experts. According to the academic-oriented evaluation approach, evaluation should be viewed as a broad process. The qualitative-oriented evaluation approach, which includes Patton's evaluation model, aims to make decisions about the curriculum, enhance its effectiveness, provide future-oriented decisions or information about the curriculum, and collect information about curriculum outcomes (Patton, 2002). The consumer-oriented evaluation approach aims to develop evaluative information about products and is based on Scriven's goal-free evaluation model (Bledsoe & Graham, 2005; Gredler, 1996).

2.2. Teachers' Role in Curriculum Assessment

Teachers play an important role in implementing educational innovations and implementing curriculum changes. Teachers' perceptions of innovations and curriculum content are very important. Teachers can take notes on the successes and failures of the curriculum implemented daily and throughout the year. Teachers can demonstrate how students' opinions about the curriculum change over time. In this way, they monitor student changes over time. Likewise, teachers can evaluate their own and their students' perceptions of the new curricula (Kelly, 2004; Ornstein & Hunkins, 2003). The best way for teachers to effectively evaluate a curriculum is to do so collaboratively. Teachers can assess the impact of the curriculum in their respective classrooms. If teachers do not collaborate and work separately, they will only see the effects of the curriculum on their own students. However, by working together, teachers can evaluate the overall impact of the curriculum. This proves that teachers are generally the most important actors in the assessment process (Taba, 1962). Another indicator of the role of teachers in curriculum evaluation is self-evaluation. Self-assessment provides important data for teachers and promotes independence rather than dependence on others (Pratt, 1994).

2.3. Kosovo Education System and Kosovo's Curriculum

Since 2000, Kosovo has undertaken important educational reforms to harmonize its system with European standards. In August 2000, a decision was made on the new 5+4+3 structure of the education system. The new 5+4+3 model replaced the existing 4+4+4 structure, that is, the new

school system extends compulsory education from 8 to 9 years, which is in accordance with the European and international flows in education. At the end of 2000 and early 2002, UNICEF and UNESCO supported the development of the new Kosovo Curriculum Framework (DASH, 2001). In the last decade, numerous Law and Bylaw Acts have been approved, many schools have been built and new educational institutions have been established. Changes have also occurred in the Curriculum Framework, as a basic education document, by which the Pre-university Education System in the Republic of Kosovo is regulated (MAShT, 2016a). Table 1 shows the levels of education. The Core Curriculum is also designed for each level.

Table 1
Pre-university Education System in the Republic of Kosovo

ISCDE & Levels	*Formal levels of the pre-university education*	*Age*	*Curriculum*
ISCDE 0	Pre-school Education	From the birth – to 5-year-old	Core Curriculum for education in the early childhood 0-5 years old
ISCDE 1			
Level 1	Pre. Class and Prim. Ed.	5-6 years old	Core Curriculum for preparatory class and primary education, grades 1-5
	Grade 1-2	6-10 years old	
Level 2	Grade 3-5		
ISCDE 2			
Level 3	Grade 6-7	11-14 years old	Core Curriculum of the Primary School (Grades, 6, 7, 8 and 9)
Level 4	Grade 8-9		
ISCDE 3			
Level 5	Grade 10-11	15-17 years old	Core Curriculum for High School (Gymnasium - Grades 10, 11 and 12) Professional School
Level 6	Grade 12		

The shift from access based on teaching objectives to access oriented learning competences and results is one of the changes from the precursory curriculum. While pedagogical currents respond to pragmatism, the curriculum's philosophy is based on progressive philosophy as the philosophy of education. Additionally, a student-centered approach to teaching and learning is promoted by the curriculum. Teachers plan and implement the interactive/comprehensive strategies. The teachers pay attention to the potential, needs and interests of the students when planning and designing lessons. The learning and teaching methods are based on the principle of inclusivity. Teachers focus on differences related to learning rhythms and styles, as well as other aspects of student diversity, including gender, age, culture, social and economic background, and students' special needs. Teachers use a variety of learning and teaching methods, techniques and materials, as well as differentiated tasks to implement the activity, aiming to positively stimulate students' interest in setting and achieving their learning outcomes (MAShT, 2016b). In addition to teaching methodology, the curriculum also promotes new assessment methods and techniques. In addition to summative assessment, teachers also use formative assessment. Based on the literature regarding teachers' role as evaluators of curriculum the research question is:

RQ) Is the curriculum assessment measure sufficiently reliable and valid from the teachers' perspective?

3. Method

3.1. Participants

The study was conducted with teachers working at various years of experiences in Kosovo. The study data were obtained from teachers who participated voluntarily. Data were collected using two different samples. Sample 1 was used for the scale items' clarity, comprehensibility, and reliability and the exploratory factor analysis process. Sample 2 was used for confirmatory factor analysis and for calculating convergent and discriminant validity. Kline (1994) has suggested that a

sample size 200 is adequate for factor analysis. However, Hair et al. (2019) recommend that the sample size for scale development studies be five times the number of items in the scale. In this context, 269 teachers constituted Sample 1 for exploratory factor analysis, while 200 teachers formed Sample 2 for confirmatory factor analysis. The demographic characteristics of the teachers are presented in Table 2.

Table 2
The demographic characteristics of the participants

Characteristics	*Sample 1*		*Sample 2*	
	n	%	*n*	%
Gender				
Female	175	65.1	141	64.1
Male	94	34.9	79	35.9
Education level				
Higher education	20	7.4	16	7.7
Undergraduate	184	68.4	127	57.7
Master's degree	59	21.9	73	32.2
Doctorate	6	2.2	3	1.4
Years of experience				
1-5 Years	70	26	49	22.3
6-10 Years	40	14.9	39	17.7
11-15 Years	40	14.9	37	16.8
16-20 Years	48	17.8	50	22.7
21 Years and Above	71	26.4	45	20.5
Employement by educational level				
Preschool	16	5.9	36	16.4
Elementary school	123	45.7	75	24.1
Middle school	110	40.9	69	31.4
High school	20	7.4	40	18.2
Receipt of in-service curriculum training				
Yes	230	85.5	188	85.4
No	39	14.5	32	14.6
Subject of teaching				
Language and Communication	18	6.7	24	10.9
Arts	12	4.5	11	5
Mathematics	20	7.4	19	8.6
Science	25	9.3	27	12.3
Social Studies and Nature	19	7.1	12	5.5
Sports and Health	17	6.3	5	2.3
Work and Life	19	7.1	11	5
Other (Preschool, Classroom)	139	51.7	93	50.5

Female teachers were higher in both Sample 1 and Sample 2. Examining the education level, most participants in Sample 1 (n=184, 68.4%) and Sample 2 (n=127, 57.7%) held undergraduate degrees. Also, the number of primary and middle school teachers was higher in both Sample 1 and Sample 2.

3.2. Data Collection Tool and Development Process

A scale has been developed intended for use in comprehensive curriculum evaluation to contribute to the literature and to pre-service and in-service teacher training. One of the objectives within this framework is to develop a scale that can be adapted to any culture. Also, it was decided that a Likert scale would be appropriate due to its structure, which allows for large-scale data collection and evaluations based on teachers' self-reported accounts.

3.2.1. Development of item pool

For the scale, literature on curriculum and change (Male, 2012; Taba, 1962; Tanner & Tanner, 1970; Van den Akker, 2003; Walker, 2003), curriculum and teacher responsibility (Adams, 2000; Clark & Yinger, 1987), curriculum and evaluation (Nouraey et al., 2020; McCormick & James, 2018; Scott, 2001), curriculum philosophy and approach (Ediger, 2003; Gutek, 1988; Heling & Bangxiu, 2009; Ornstein, 1990), and curriculum implementation (Adams, 2000; Evans, 2003; Karakus, 2021; Lewthwaite, 2001; Nevenglosky, 2018; VanTassel-Baska & Brown, 2007; Van den Akker et al., 2006) was reviewed. Items were prepared to cover program approach and philosophy, planning, learning outcomes, content, learning-teaching process, assessment and evaluation, resources materials (books), and curriculum implementation. These items were associated with principles such as coherence between curriculum elements, teacher competence, duration, vertical and horizontal integration of topics, teacher training, teacher competence, student relevance (suitability to readiness), alignment with societal expectations, clarity, comprehensibility, currency, modernity, physical facilities, and support. The item pool was prepared in a two-dimensional format, as shown in Table 3.

Table 3
The Item Pool Matrix

Aspects / Qualities	*Cl*	*R*	*A*	*SO*	*Co*	*F*	*D*	*TCT*	*CU*	*App*
Curriculum philosophy/approach	X		X					X	X	
Content		X	X	X	X			X	X	
Measurement and evaluation			X					X		X
The process of teaching and learning	X		X				X	X		X
Learning outcomes	X			X					X	
Curriculum tools (books, materials, etc.)		X	X			X			X	
Planning				X			X	X		X

Note. Cl: Clarity; R: Relevance; A: Association; SO: Student-orientedness; Co: Coherence; D: Duration; TCT: Teacher Competency-Training; CU: Clarity and Understandability; App: Applicability.

Within the framework of the matrix mentioned above, a pool of 36 items was created. The prepared item pool was discussed face-to-face with five experts in the field of curriculum development. The principles proposed by Harrison and McLaughlin (1993) were considered. The experts were asked to examine the matrix in Table 3 for its suitability, as well as the clarity and comprehensibility of the expressions. Based on the experts' recommendations, some expressions (3 items) were modified regarding wording and phrasing. For example, the item "I need additional training to implement the curriculum most effectively" was changed to "I need training (in-service training, workshops, etc.) to implement the curriculum effectively."

3.2.2. Evaluation of content adequacy of items

The suitability of the items in terms of content has been ensured. Nunnally and Bernstein (1994) emphasized the importance of eliminating inconsistent and irrelevant items and adding any missing ones before collecting and analysing data for the pilot test of a scale. In this context, the expert group was asked whether there were any irrelevant existing items or if any additional items should be included. The experts did not identify any irrelevant or insufficient items. Additionally, they indicated that all items conformed to the Item Pool Matrix.

For further evaluation of content adequacy, three teachers and one prospective teacher from the target group were asked to read the scale items separately but face-to-face. These teachers were then asked to group the items that could be related to each other. In the independent groupings made by each of the three participants, one teacher evaluated the item stated as "Teaching and learning methods ensure students learn by doing and experiencing" within multiple groups. Furthermore, it was determined that all items were similarly grouped under a specific category.

Also, teachers were asked to evaluate the readability, clarity, and comprehensibility of each item's intended content. The teachers stated that the characteristic intended to be measured by each item was clear and understandable.

3.2.3. Preparation and application of the scale

The prepared 36 items were converted into a 5-point Likert scale: 1 = Strongly Disagree, 2 = Disagree, 3 = Neutral, 4 = Agree, and 5 = Strongly Agree. The 5-point Likert form was first tested for trial application for Exploratory Factor Analysis [EFA] and then prepared for Confirmatory Factor Analysis [CFA] following the EFA. The application for EFA and CFA was conducted face-to-face in schools during the spring semester of the 2022-2023 academic year. The purpose of the study was explained to the teachers and administered to those who voluntarily agreed to participate. Teachers sampled for the EFA analysis were not included in the CFA analysis. Therefore, the CFA was conducted in schools different from those where the EFA sample was collected, involving other volunteer teachers.

3.3. Data Analysis

To perform the measurement tool's validity and reliability analysis, the data obtained from Sample 1 and Sample 2 were analysed using SPSS 23.00 for Descriptive Analysis and EFA, Jamovi 2.3.18 and JASP 0.9.0.1 for Multivariate Normality, Alpha, and OMEGA reliability analyses, and Lisrel 8.7 for CFA.

To determine the factor structure of the scale, EFA based on Maximum Likelihood (ML) was conducted using data from Sample 1 (Colton & Covert, 2007; Comrey & Lee, 1992). ML requires that the data exhibit a normal distribution (Fabrigar et al., 1999; MacCallum et al., 2007). Therefore, multivariate normality was tested first. This test was selected because of its general applicability to normally distributed data, preference for more significant correlations, and lower variability of estimates compared to other models (Briggs & MacCallum, 2003; Fabrigar et al.,1999). Factor analysis aims to combine many interrelated measurements into typical structures or factors. Since factor analysis assumes that all variables are somewhat correlated (Byrne, 2001; Kandemir et al., 2019), this analysis was performed to identify unrelated items that did not fall under any factor. ML determined the minimum number of factors appropriate for the original data set (Ford et al., 1986). The number of factors was determined using eigenvalues (significant for factors with an eigenvalue of 1 or higher), the Scree Test, and Velicer's Minimum Average Partial [MAP] Test (Hair et al., 2019; Velicer, 1976; Velicer et al., 2000). Additionally, the "direct oblimin" technique, one of the oblique rotation techniques, was used. This technique is preferred because dimensions in behavioural science fields are expected to be correlated (Browne, 2001; Williams et al., 2020). Tabachnick and Fidell (2001) suggested that oblique rotation should be preferred when there is no significant justification and if the correlation matrix contains coefficients of 0.32 or higher. Since factor loadings met practical significance, a value of ± 0.30 was used. This value was chosen because it contributes to explaining the amount of variance accounted for by a factor (Ho, 2006).

CFA was conducted using data from Sample 2. CFA was used to test the accuracy of the structure determined by EFA. CFA was conducted using the Maximum Likelihood method in the Lisrel program. It was applied to test the factorial structure of the model determined by EFA (Ding et al., 1995; Gomez & Fisher, 2003). A series of indices were used to evaluate the model's fit. The first of these, the x^2 index, is assessed together with the degrees of freedom due to its sensitivity to sample size. Also, the CFI (Comparative Fit Index), GFI (Goodness of Fit Index), NNFI (Non-Normed Fit Index), and NFI (Normed Fit Index) values are considered acceptable if they are close to 1. However, values of 0.90 or higher are also acceptable (Bentler & Bonett, 1980). Hu and Bentler (1999) noted that values of 0.95 or higher indicate a good fit. RMSEA (Root Mean Square Error of Approximation) values of 0.08 or lower are acceptable, with 0.06 indicating a better fit (Bentler & Bonett, 1980).

The convergent validity of the scale was determined by analysing the Explained Common Variance [ECV] values for each factor and comparing the correlations among the factors (Fornell &

Larcker, 1981). Discriminant validity was evaluated by comparing the square root of the variance explained with the square of the correlations among factors. Convergent and discriminant validity are other types used in testing and confirming the model (Fornell & Larcker, 1981; Malhotra, 2011). Cronbach's Alpha, Omega, and Composite Reliability were calculated for reliability analysis. Composite Reliability [CR], used to measure the internal consistency of factors, is considered good if it is 0.70 or higher (Hair et al., 2019). In the context of internal consistency, Cronbach's Alpha analysis alone is deemed insufficient for multifactorial structures. It is also recommended that the Omega Reliability coefficient be calculated (Dunn et al., 2014).

4. Results

4.1. Suitability of Data for Analysis

To determine the suitability of the scale data for factor analysis, the Kaiser-Meyer-Olkin [KMO] test and Bartlett's test were conducted. The KMO test evaluates the adequacy of each observed value for sampling. It is calculated based on the correlations among variables. A KMO value of 0.70 or above is considered sufficient (Kaiser, 1974). Bartlett's test of sphericity tests the hypothesis that the correlation matrix is an identity matrix. A significant result indicates that the correlation matrix is suitable for factor analysis (Bartlett, 1951). In this study, the KMO measure of sampling adequacy was found to be 0.752, indicating sufficient sampling adequacy. Bartlett's test of sphericity yielded a value of Approx. Chi-Square = 3922.520 (df = 741, $p < .05$), which was significant, indicating that the correlation matrix is not an identity matrix.

To determine the factor structure of the measurement instrument, all data were observed to be normally distributed, and no outliers were detected. The data obtained from the application for Exploratory Factor Analysis and the results of the Descriptive Analysis are presented in Table 4.

Table 4
Descriptive Statistics for Sample 1

Items	*M*	*SD*	*Skw*	*Krt.*
M1. The high number of students makes it difficult for me to implement the curriculum.	3.36	1.22	−.446	−1.022
M2. The curriculum philosophy centers around the student.	3.48	.90	−.834	.398
M3. The curriculum philosophy contradicts my educational beliefs.	3.88	1.02	−.206	.873
M4. Learning outcomes meet the expectations of society.	3.07	1.24	−.269	−.925
M5. The content (concepts and learning topics) aims to achieve the course's learning outcomes.	3.49	1.19	−.717	−.521
M6. The learning outcomes are different from the goals set in the curriculum.	3.11	1.16	−.376	−.793
M7. The content (concepts and learning topics) is aimed at achieving competencies.	3.18	1.09	−.781	−.672
M8. We collaborate with teachers from other courses and disciplines to prepare plans.	2.60	1.12	.439	−.812
M9. Teaching and learning methods facilitate students' learning by doing and experiencing.	3.18	1.22	−.243	−1.101
M10. Throughout the lesson planning, connections are made with subject area topics.	3.72	1.08	−.862	−.074
M11. The assessment guide is complex.	3.25	1.12	−.503	−.697
M12. The teaching units in the textbooks are up to date.	2.60	1.18	.360	−.980
M13. Learning outcomes for competencies are understandable.	3.26	1.14	−.434	−.666
M14. Although the curriculum is well-written, it is difficult to implement in practice.	3.00	1.23	.070	−1.183

Table 4 continued

Items	*M*	*SD*	*Skw*	*Krt.*
M15. The methodologies presented in the curriculum could be more practical.	2.90	1.18	.127	−1.110
M16. The content (concepts and learning topics) is relevant to daily life.	3.30	1.09	−.347	−.925
M17. The content (concepts and learning topics) is updated.	3.15	1.10	−.169	−1.06
M18. The books do not align with the curriculum content (concepts, topics, etc.).	3.15	1.19	−.424	−.985
M19. Contemporary assessment approaches are also anticipated (portfolio, self-assessment, performance assessment, etc.).	3.85	1.07	−1.096	.726
M20. The assessment is consistent with the learning outcomes.	3.43	1.05	−.740	−.354
M21. Learning outcomes for competencies are consistent with the student level.	3.23	1.13	−.278	−.913
M22. Assessment supports the development of self-assessment skills in students.	3.36	1.07	−.604	−.604
M23. The content (concepts and learning topics) suits the student's level.	3.21	1.08	−.271	−1.126
M24. The teaching units in the books align with the envisaged concepts and topics in the curriculum.	2.88	1.15	−.014	−1.200
M25. I lack sufficient knowledge about the use of teaching tools and materials.	3.03	1.16	−.123	−1.268
M26. The assessment tools provided for assessment are adequate.	3.66	1.03	−.809	.090
M27. Sufficient tools and materials are available at the school to implement the curriculum.	2.34	1.11	.800	−.206
M28. I have sufficient knowledge about the methodologies presented in the curriculum.	3.15	1.19	−.424	−.985
M29. There is no need for planning the curriculum.	3.27	1.26	−.163	−1.28
M30. The plan presented in the curriculum needs to be functional.	3.00	1.28	.145	−1.23
M31. The numerous topics included in the curriculum's content make its implementation easier.	3.02	1.22	.047	−.680
M32. Throughout the lesson planning, connections are made with topics from other fields.	3.96	.87	−.670	1.116
M33. I need help to work with teachers from other courses and disciplines.	2.40	1.30	.767	−.612
M34. I need training (in-service training, workshops, etc.) to implement the curriculum effectively.	3.52	1.10	−.753	−.268
M35. The current teaching hours are not sufficient for implementing the curriculum.	3.33	1.12	−.397	−.872
M36. The methodologies presented in the curriculum are compatible with the teaching and learning process.	2.58	1.18	.503	−.820

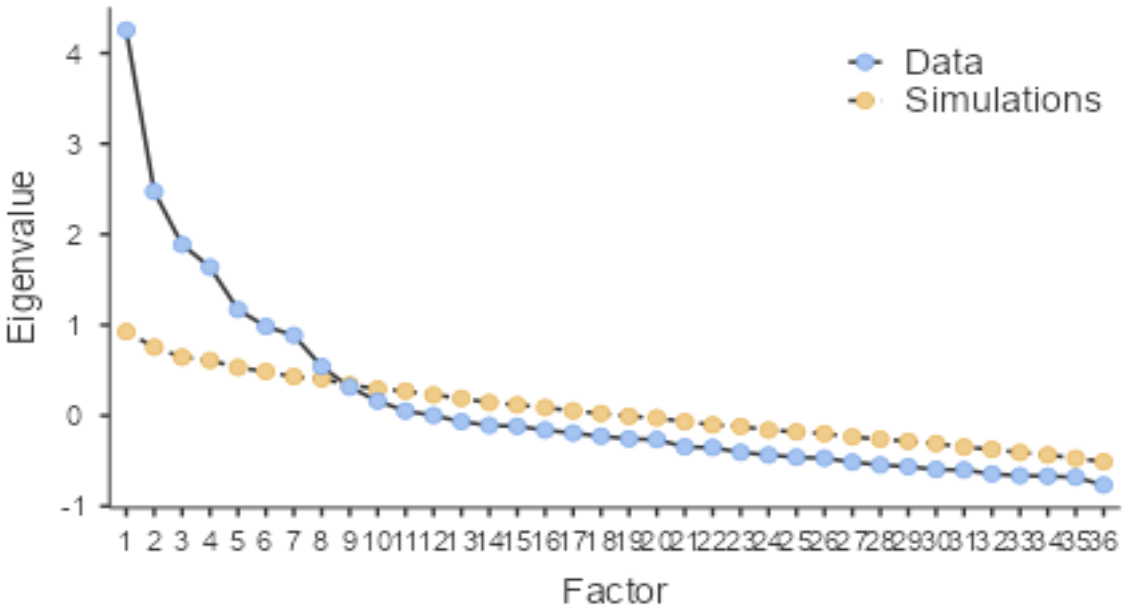
According to the analysis results, the item "I relate course topics to subject area topics during lesson planning" had the highest mean score (M=3.96, SD=.87). In contrast, the item "I struggle to collaborate with teachers in other courses and areas" had the lowest mean score (M=2.20, SD=1.11). The skewness and kurtosis values for all items were within the ±2 range. P-P and Q-Q plots were examined to assess the normality of all items. Mahalanobis D^2 distances were analysed to identify potential outliers (Tabachnick & Fidell, 2001). Only one outlier was detected and subsequently removed from the data set. The data were observed to conform to the assumptions of multivariate normality. Additionally, bivariate and partial correlations among the items were examined. The highest bivariate correlation (r = .638, p < .05) was found between Item 8 and Item 30, while the lowest correlation (r = .039, p>.05) was observed between Item 1 and Item 6. The highest bivariate correlation for Item 1 was with Item 35 (r = .43, p < .05). The bivariate correlations among the other

items were below these values. The multicollinearity among the variables was tested using the Variance Inflation Factor [VIF], and no multicollinearity was observed.

4.2. Analysis of Factor Structure

Determining the factor structure of a scale can be achieved through various methods. One of the most commonly used techniques is the Scree test, as proposed by Cattell (1978). However, this graphical method has been criticized for its reliance on visual interpretation. Furthermore, opinions suggest this method is more suitable for large samples (Zwick & Velicer, 1986). Another approach is to select factors with eigenvalues greater than 1 (Kline, 1994; Tabachnick & Fidell, 2001). This method has also been evaluated for potentially including some sampling errors and producing more factors (Thompson, 2004; Zwick & Velicer, 1986). Another method is the Minimum Average Partial [MAP] test proposed by Velicer (1976). In this study, all three approaches have been examined. As illustrated in Figure 1, the Scree Test indicates that there may be 8 factors for the scale's 36-item structure.

Figure 1
Scree plot



Although there were 10 factors with eigenvalues greater than 1, the ninth factor contained 2 items, and the tenth factor contained only 1 item. Maccallum et al., (1999) state that for a factor to be considered valid, it should have at least 3 items. When these factors are disregarded, an 8-factor structure is more appropriate. It has been reported that parallel analysis (Horn, 1965) and the MAP test (Velicer, 1976) produce similar results (Ladesma & Valero-Mora, 2007; O'Connor, 2000; Yavuz & Doğan, 2015; Zwick & Velicer, 1986). In this study, the procedure recommended by O'Connor (2000) was used to determine the number of factors. The results of the Average Partial Correlation analysis are presented in Table 5.

The analysis of the MAP test results indicated that the smallest mean squares partial correlation was 0.0154, observed at the 8th step. It was determined that the fourth power of the partial correlation occurred at the 8th step. The fourth power of the partial correlation was included in the program by O'Conner (2000). Considering the scree test, eigenvalues, and MAP analysis, it was assessed that the number of factors would be 8.

Table 5
Eigenvalues Regarding Partial Correlations Obtained from the MAP Test

	Squared	*Power4*		*Squared*	*Power4*		*Squared*	*Power4*
0	.0298	.0048	12	.0209	.0016	24	.0688	.0146
1	.0231	.0028	13	.0228	.0022	25	.0771	.0184
2	.0201	.0016	14	.0250	.0027	26	.0868	.0232
3	.0194	.0012	15	.0267	.0030	27	.0997	.0276
4	.0181	.0010	16	.0293	.0034	28	.1125	.0339
5	.0170	.0009	17	.0324	.0048	29	.1304	.0432
6	.0163	.0007	18	.0357	.0051	30	.1517	.0567
7	.0157	.0007	19	.0394	.0059	31	.1854	.0741
8	.0154	.0007	20	.0440	.0070	32	.2458	.1196
9	.0169	.0012	21	.0491	.0085	33	.3280	.1955
10	.0181	.0013	22	.0550	.0098	34	.4877	.3616
11	.0194	.0015	23	.0616	.0119	35	1.0000	1.0000

4.3. Exploratory Factor Analysis

Data from 269 teachers, consisting of 36 items, were analysed using EFA with the Maximum Likelihood method. Fabrigar et al. (1999) pointed out that when a scale structure contains multiple factors and there is a correlation among these factors, oblique rotations are significant for identifying accurate and realistic factors. Therefore, the Direct Oblimin method was employed for rotation. The factors obtained from the analysis and the items within these factors are presented in Table 6.

Table 6
The Distribution of Items to the Factors

	Structure Matrix									
Item No	1	2	3	4	5	6	7	8	9	10
M34*	.963						.250		.287	
M3*	.312		-.232							
M22		.851					.259	.356		.205
M11		.734			.230			.267		.240
M26		.721					.248	.335		
M19		.690					.282	.250		
M20		.557					.225			
M16			.835	−.214						
M5			.813							
M23			.695							
M17			.672							
M30			.252	.912						
M8			.230	.681						
M29				.567	.222					
M9					.769					.254
M2					.681					
M10					.677					
M32					.585					
M33*		.208		−.234	.351		.347	.209		
M15						.845				
M36						.653				
M28						.549				
M21		.309					.796	.230		
M4		.298					.790	.209		

Table 6 continued

	Structure Matrix									
Item No	1	2	3	4	5	6	7	8	9	10
M13		.251					.751			−258
M6							.529		.212	−.287
M27		.336						.824		
M12		.221						.732		
M24		.235						.654		
M25		.239						.625		
M31									.772	
M35									.643	
M1									.622	
M14	.220								.566	
M18*										.254
M7*					.208					.219
Percentile Variance Explained for 10 Factors	4.601	12.067	7.505	5.574	5.491	4.426	4.257	3.613	2.744	1.401
Percentile Variance Explained for 8 Factors	13.837	9.095	6.236	6.623	5.120	4.474	4.056	2.808		

Note. *Items omitted in the second analysis.

As a result of the analysis based on Direct Oblimin rotation, items 34 and 3 formed a single factor. However, item 3, like item 33, exhibited cross-loading, meaning that its factor loadings were below 0.100 under multiple factors; therefore, it was excluded from the scale. In this context, item 34 was considered a standalone factor, and items 7 and 18 were excluded from the scale because their factor loadings were below 0.30. Consequently, the factor structure was reduced from 10 factors with eigenvalues greater than 1 to an 8-factor structure. The explained variance increased from 51.680% to 52.550%. The Velicer's MAP test (Velicer, 1976; Velicer et al., 2000) and the Scree Plot (Figure 1) similarly indicated an 8-factor structure. The Goodness-of-fit Test yielded a significant Chi-Square value of 357.124 ($df= 245$, $p<.05$).

Each factor was named by examining the items within each factor. Accordingly, the factors were labeled as follows: the first-factor "Assessment and Evaluation" the second-factor "Content" the third-factor "Learning Outcomes," the fourth-factor "Resources/Materials," the fifth-factor "Planning," the sixth-factor "Curriculum Implementation," the seventh-factor "Curriculum Approach/Philosophy," and the eighth-factor "Methodology."

4.4. Confirmatory Factor Analysis Results

To verify the factors obtained from the exploratory factor analysis, a CFA was conducted. This analysis was performed on data collected from a new sample group, distinct from the one used in the EFA. Prior to the analysis, the distribution of the data was examined. The multivariate normality of the data obtained from Sample 2 was tested by assessing Skewness and Kurtosis values, checking for outliers using P-P and Q-Q plots, and analysing Histograms and Mahalanobis D^2 distances. The descriptive analysis results of the data from Sample 2 are presented in Table 7.

Table 7
Descriptive Analysis for Sample 2

		M	*SD*	*Skewness*	*Kurtosis*
AE11	The assessment guide is complicated.	3.26	1.18	−.350	−.842
AE19	Contemporary approaches to assessment are also envisaged (such as portfolio, self-assessment, performance appraisal, etc.).	3.23	1.13	−.381	−.664
AE20	Assessment is consistent with learning outcomes.	3.22	1.13	−.372	−.687
AE22	Assessment supports the development of students' self-assessment skills.	3.20	1.11	−.252	−.743
AE26	The assessment instruments provided for assessment are adequate.	3.18	1.13	−.325	−.752
	Assessment and Evaluation	3.22	.99	−.315	−.693
CO5	The content (concepts and learning topics) aims to achieve the course's learning outcomes.	3.33	1.26	−.371	−.902
CO16	The content (concepts and learning topics) is relevant to daily life.	3.38	1.27	−.345	−.938
CO17	The content (concepts and learning topics) is updated.	3.26	1.23	−.244	−.879
CO23	The content (concepts and learning topics) suits the students' level.	3.20	1.21	−.220	−.872
Content		3.31	1.16	−.249	−1.020
LO4	Learning outcomes meet the expectations of society.	3.25	1.07	−.402	−.698
LO6	Learning outcomes are not compatible with the objectives set in the curriculum.	3.24	1.14	−.512	−.844
LO13	Learning outcomes for competencies are understandable.	3.09	1.19	−.250	−1.118
LO21	Learning outcomes for competencies are consistent with the student level.	3.08	1.16	−.099	−1.103
Learning Output		3.16	.90	−.338	−.389
IS12	The teaching units in the textbooks are current.	3.09	1.25	−.098	−1.211
IS24	The teaching units in the textbooks suit the concepts and topics envisaged in the curriculum.	3.22	1.20	−.166	−1.122
IS25	I lack sufficient knowledge about the use of teaching tools and materials.	3.20	1.20	−.298	−1.070
IS27	The school has sufficient teaching tools and materials to implement the curriculum.	3.28	1.20	−.425	−.915
Resources/Instruments		3.20	.95	−.383	−.437
PL8	We prepare plans together with teachers from other courses and fields.	3.04	1.20	−.329	−1.092
PL29	There is no need for planning the curriculum.	3.11	1.13	−.132	−.945
PL30	The plan presented in the curriculum needs to be functional.	3.13	1.21	−.373	−.859
Planning		3.09	.91	−.448	−.371
AP1	The abundance of students poses a challenge to the implementation of the curriculum.	3.31	1.18	−.623	−.605
AP14	Despite the curriculum being well-designed in theory, its practical implementation presents difficulties.	3.20	1.11	−.392	−.675
AP31	The multitude of topics covered in the curriculum's content complicates its implementation.	2.75	1.17	.446	−.779
AP35	The current teaching hours need to be increased for the effective execution of the curriculum.	3.13	1.20	−.225	−1.009

Table 7 continued

		M	*SD*	*Skewness*	*Kurtosis*
Curriculum Application		3.10	.88	−.333	.033
CP2	The philosophy of the curriculum places the student at the centre.	2.87	1.19	.344	−.763
CP9	Teaching and learning methods enable students to learn by doing and experiencing.	2.85	1.23	.372	−1.008
CP10	Throughout the course planning, connections are established with subject matters within the field.	2.87	1.23	.132	−1.063
CP32	Throughout the course planning, connections are established with subject matters in other fields.	2.63	1.32	.480	−.925
Curriculum Philosophy		2.81	.99	.407	−.418
MT15	The methodologies presented in the curriculum need to be more applicable.	2.55	1.23	.506	−.780
MT28	I am knowledgeable about the methodologies presented in the curriculum.	3.25	1.16	−.351	−.716
MT36	The methodologies presented in the curriculum are compatible with the teaching and learning process.	2.76	1.22	.225	−1.091
Teaching Methods		2.86	.94	.067	−.287
Overall Mean		3.09	.94	−.372	.549

According to the descriptive analysis results of Sample 2, the highest mean (M=3.31, SD=1.16) was observed in the Content dimension, while the lowest mean (M=2.81, SD=0.99) was observed in the Curriculum Approach/Philosophy dimension. To test the accuracy of the factor structures identified through the exploratory factor analysis of the scale, a Confirmatory Factor Analysis was applied to Sample 2. Since the second dataset of the scale also demonstrated multivariate normality, the Maximum Likelihood method, a robust estimation technique, was used. The fit indices obtained from the analysis are presented in Table 8.

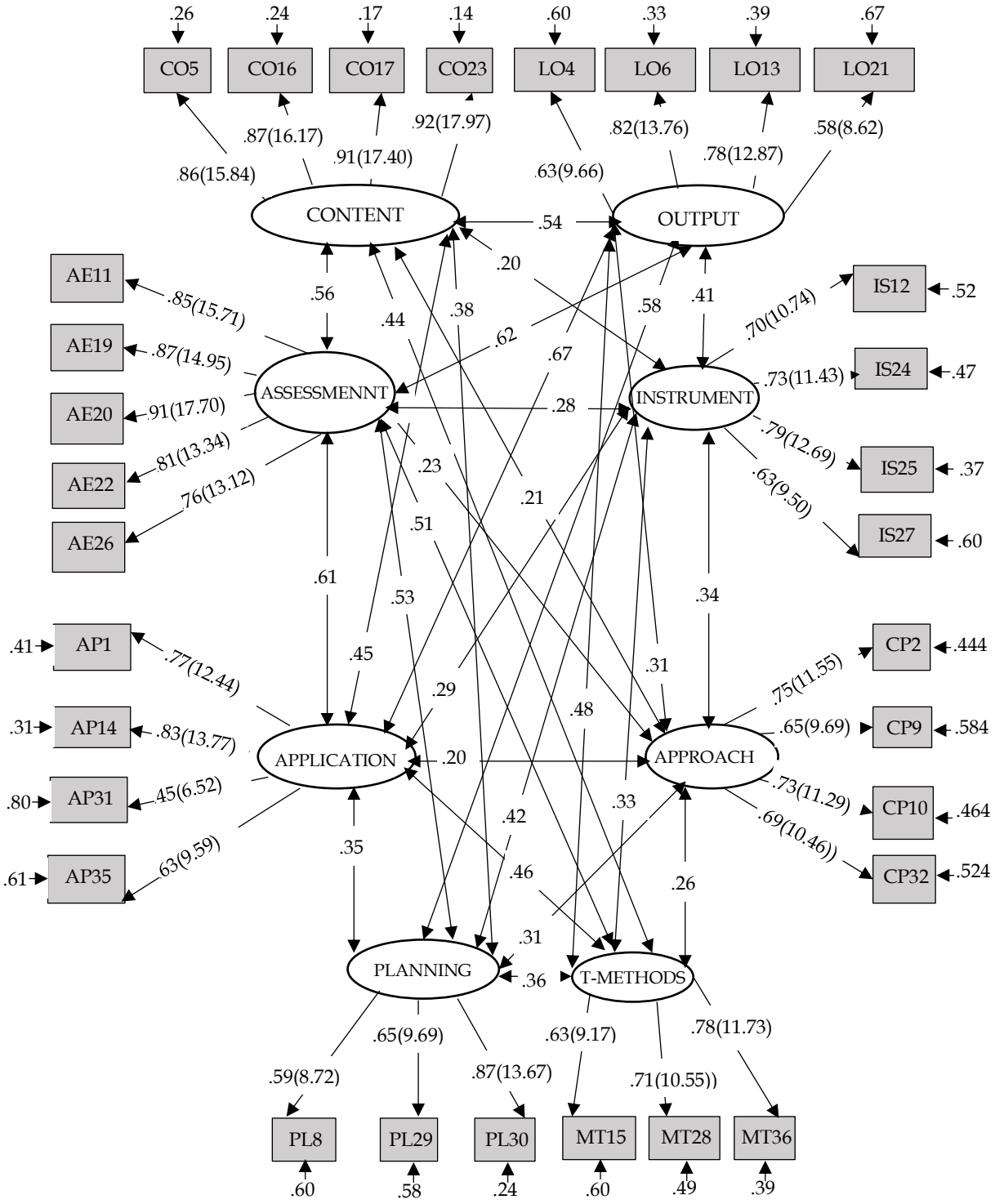
Table 8
Fit Indices according to the CFA

Index	*Chi-Square*	*Df*	*RMSEA*	*NFI*	*NNFI*	*CFI*	*IFI*	*RMR*	*SRMR*	*GFI*	*AGFI*
Value	666.07	406	.054	.93	.96	.97	.97	.081	.057	.84	.80
Modification AE19-AE22 and LO21-Lo4	632.55	404	.051	.93	.97	.97	.97	.080	.056	.84	.81

According to the analysis results, the Chi-Square (665.35) to degrees of freedom (405) ratio was 1.64. Additionally, the RMSEA value (.056) is within an acceptable range. The NNFI, CFI, and IFI indices are excellent, the NFI is good, and the RMR, SRMR, GFI, and AGFI indices are within acceptable limits. The correction based on two modifications (AE19 and AE22; LO21 and LO4) in the error variances did not sufficiently increase the GFI and AGFI indices. The path coefficients and *t*-values resulting from the analysis are presented in Figure 2.

In the DFA results, the paths drawn from observed variables to latent variables were found to be significant ($p < .05$). The item with the lowest path coefficient ($\lambda = .45$, $t = 6.52$, $p < .05$) belonged to the Curriculum Implementation dimension, specifically the AP31 item, which is expressed as "The large number of topics in the curriculum content makes it difficult to implement the curriculum." The item with the highest path coefficient ($\lambda=.92$, $t = 17.97$, $p < .05$) belonged to the Content dimension, specifically the CO23 item, which is expressed as "The content (concepts and learning topics) is appropriate for the student's level." Significant correlations were also obtained between the latent variables. The lowest correlation was observed between Curriculum Approach/Philosophy and Teaching Methods at 0.17 ($t = 1.98$, $p < .05$), and the highest correlation was observed between Planning and Learning Output at 0.69 ($t = 13.34$, $p < .05$).

Figure 2
CFA Path Diagram



4.5. Convergence and Discriminant Validity

Although CFA is used for construct validity, Campbell and Fiske (1959) suggested examining convergent and discriminant validity to determine the structure of a measurement instrument. Convergent validity refers to the degree of confidence that the indicators measure the intended construct. Discriminant validity, on the other hand, refers to the extent to which different constructs that are not supposed to be related are indeed unrelated. Discriminant validity determines whether the observed variables represent the underlying latent structures they are associated with (Hair et al., 2019). The convergent validity of the measurement model is examined

through the values of AVE and CR. An acceptable CR value is 0.70 or higher, while an acceptable AVE value is ideally 0.70 or higher, but 0.50 or higher is sufficient. The CR value should also be greater than the AVE value (Gouveia & Soares, 2015; Raykov, 1997). Moreover, the square root of the AVE value should be greater than the correlation values between the latent variables (Hu & Bentler, 1999). The Maximum Shared Variance [MSV] and Average Shared Variance [ASV] values are examined for discriminant validity. Hair et al. (2019) recommend the criteria AVE > MSV and AVE > ASV for evaluating discriminant validity. The results of the analyses conducted within this framework are presented in Table 9.

Table 9
CR, AVE MSV, ASV, and Correlations between Dimensions

	CR	*AVE*	*MSV*	*ASV*	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*
1.Assessment and Evaluation	.92	.71	.38	.24	(.84)							
2. Content	.94	.79	.31	.17	.56	(.89)						
3. Learning Output	.81	.52	.44	.28	.62	.54	(.72)					
4.Instrument/Resources	.81	.52	.18	.11	.28	.20	.41	(.72)				
5.Planning	.75	.51	.33	.20	.53	.38	.57	.42	(.71)			
6.Curriculum Implementation	.77	.48	.44	.21	.61	.45	.66	.29	.35	(.69)		
7.Curriculum Approaches/Philosophy	.80	.50	.12	.08	.22	.21	.32	.34	.31	.20	(.71)	
8.Teaching Methods	.75	.50	.25	.17	.57	.44	.48	.33	.36	.46	.26	(.71)

Note. Square roots of average variances extracted are shown on a diagonal.

According to Fornell and Larcker (1981), a CR value above 0.60 and an AVE value above 0.50 are sufficient for establishing discriminant validity. However, when the CR value is 0.70 or higher, an AVE value of 0.40 or higher is considered sufficient for discriminant validity. In this study, the AVE value for the Curriculum Implementation dimension was 0.48. Nevertheless, since the CR value was 0.77, it was deemed sufficient for discriminant validity. For the other dimensions, the AVE values were 0.50 or higher, and the CR values were above 0.70. Additionally, it was observed that AVE > MSV and ASV (Hair et al., 2019). The results indicated that the scale possesses both convergent and discriminant validity.

4.6. Correlations among Scale Dimensions

A Pearson moment correlation analysis was conducted to determine the correlations among the scale's factors. The results are presented in Table 10.

Table 10
Correlations Among Dimensions

	AE	*C*	*LO*	*IR*	*P*	*A*	*AP*
Content	.384*						
Learning Output	.433*	.492*					
Instrument/Resources	.290*	.278*	.419*				
Planning	.163*	.389*	.348*	.526*			
Application	.280*	.431*	.409*	.372*	.302*		
Approach/Philosophy	.205	.229*	.312*	.309*	.519*	.285*	
Teaching Methods	.241*	.332*	.293*	.427*	.318*	.440*	.393*

Note. AE: Assessment and evaluation; C: Content; LO: Learning output; IR: Instrument/ Resources ; A: Application; AP: Approach/ Philosophy; *$p < .05$.

The results of the correlation analysis indicate that the lowest correlation ($r = .16$, $p < .05$) was observed between the "Planning" and "Assessment and Evaluation" dimensions. The highest correlation ($r = .53$, $p < .05$) was identified between "Planning" and "Instrument/Resources." The

relationship between "Assessment and Evaluation" and "Planning" is weak. Correlations between the other dimensions are moderate or close to moderate.

4.7. The Reliability and Item Discrimination of the Scale

Table 11 presents the results of the analysis for each subdimension and the overall scale regarding Cronbach's Alpha reliability, item-total correlation, and the top and bottom 27% item discrimination index.

Table 11
Item Total Correlations and Reliability Coefficients

Factor Name	*Items*	*Item Total Correlation*	*Cronbach Alpha*	*Omega*
Assessment and Evaluation	AE11	.826	.92	.92
	AE19	.832		
	AE20	.871		
	AE22	.762		
	AE26	.714		
Content	CO5	.836	.94	.94
	CO16	.853		
	CO17	.869		
	CO23	.894		
Learning Output	LO4	.595	.80	.79
	LO6	.665		
	LO13	.628		
	LO21	.544		
Teaching Methods	MT15	.542	.74	.75
	MT28	.534		
	MT36	.633		
Planning	PL8	.544	.76	.77
	PL29	.579		
	PL30	.655		
Curriculum Application	AP1	.643	.75	.77
	AP14	.676		
	AP31	.380		
	AP35	.515		
Curriculum Approach Philosophy	CP2	.652	.80	.80
	CP9	.558		
	CP10	.625		
	CP32	.631		
Instrument/ Resources	IS12	.590	.80	.80
	IS24	.611		
	IS25	.688		
	IS27	.543		

As a result of the reliability analysis, the highest Cronbach's Alpha reliability coefficient was observed in the Content Dimension, with a value of 0.94. Similarly, the Omega reliability coefficient was also highest in the Content Dimension, with a value of 0.94. The lowest Cronbach's Alpha reliability was observed in the Teaching Method Dimension, with a value of 0.74, and the Omega reliability coefficient in this dimension was also observed to be 0.75. The item discrimination was analysed using the 27% upper-lower group technique. The analysis revealed significant differences between the upper and lower groups for all items, each sub-dimension, and the overall total. The *t*-values of the items were determined, with the lowest being 4.794 ($p < .05$) for item 31 and the highest being 7.473 ($p < .05$).

4.8. Analysis of In-Service Training Status of Teachers Regarding the Curriculum

In the context of educational curriculum development in Kosovo, while a group of teachers has received in-service training regarding the renewed curriculum, some have yet to participate in in-service activities during the execution of this study. In-service training has been evaluated as a criterion for obtaining in-depth knowledge about curriculum and being literate about the existing curriculum. Within this framework, the views of teachers who participated and did not participate in in-service training on curriculum evaluation have been analysed. As a result of the analysis,

significant differences have been identified in all dimensions. The highest difference (t = 14.061, p < .05) was observed in the Assessment and Evaluation dimension, while the lowest difference (t = 3.765, p < .05) was observed in the Curriculum Approach/Philosophy dimension. The effect size in the Curriculum Approach/Philosophy dimension (Cohen d = .61) is of moderate magnitude, while in the other dimensions, it has been observed that all have effect sizes above 0.80.

5. Discussion

This study examined the adequacy of the scale prepared for curriculum assessment in terms of its validity and reliability. The scale development was based on teachers' subjective self-reported assessment. Scale items were prepared considering the structure and implementation of the curriculum, as well as the qualities a curriculum should possess. Within this framework, the conformity of scale items to criteria determined in a two-dimensional matrix was checked. In this matrix, one of the dimensions includes curriculum elements, while the other dimension expresses the qualifications that curriculum elements should have. A total of 36 items were prepared in this context. Subsequently, the prepared items were presented to experts. After teachers and experts evaluated the suitability, clarity, intelligibility, comprehensiveness of the items, the final form was given. The feedback obtained from experts and teachers shows that the items are clear, comprehensible and sufficient in terms of scope.

Before the analysis, it was examined for which analysis the data obtained from the first sample was more suitable, the sample adequacy and normality, and whether there were any outliers. The analysis results showed that the number of samples was suitable for analysis, the data showed multi-variety normality, and there were no outliers. For this reason, it was decided to conduct factor analysis using the maximum likelihood method, which is a stronger estimation method in cases where a normal distribution is appropriate (Fabrigar et al., 1999). For the reliability and validity analysis of the scale, it was administered to 269 teachers. EFA based on the Maximum Likelihood method was conducted for construct validity analysis. Before conducting the validity analysis, Map Test, parallel analysis, scree test and eigenvalue were examined for the number of factors of the scale. The analysis results to determine the number of factors showed that the 8-factor structure was appropriate. As a result of exploratory factor analysis, 5 items were removed from the scale. It was observed that these items were related to curriculum approach/philosophy, curriculum implementation, and instruments/resources. These items were found to maintain the scope of the prepared scale. When evaluating the factor structure of the scale, it was considered that each factor should have at least three items, the factor loading should be above 0.30, and the loading of an item with multiple factors should be at least 0.100 greater than the loading of the same item in other factors (Maccallum et al., 1999). As a result of exploratory factor analysis, 8-factor structures were determined. The factors were named assessment and evaluation, content, learning output, teaching methods, curriculum planning, curriculum approach/philosophy, curriculum application, and instruments/resources, considering the characteristics measured by the items.

It was evaluated that two items expressed in the dimension of curriculum approach/philosophy, "Item 10: Connections have been established with topics in the field throughout lesson planning" and "Item 32: Connections have been established with topics in other fields throughout lesson planning," were related to both the curriculum structure and the curriculum approach. Therefore, this dimension was expressed with two concepts. Some researchers (Heyman, 1981; Posner & Strike, 1974) have addressed the relationship between curriculum approach and curriculum structure.

After determining the factor structures of the scale through EFA, CFA was conducted by administering the 31-item scale to 220 teachers who were not included in the initial sample. The ML estimation method was also applied in CFA. As a result of the analysis, it was observed that the path coefficients ranged from 0.45 to 0.92, indicating significant relationships between all latent

variables. Moreover, it was determined that the fit indices of the model were satisfactory. Additionally, convergence and discriminant validity analyses were conducted to determine the structural validity of the scale. The analysis results indicated satisfactory results for the convergence and discriminant validity of the scale. Moderate correlations were found between the dimensions of the scale. Finally, it was observed that the item-total correlations and reliability analysis results were high.

The results produce sufficient and reliable outcomes regarding the construct validity, content validity, convergence, and discriminant validity of the Curriculum Assessment scale. The research was conducted with teachers in Kosovo. However, testing the scale's applicability in different cultures will contribute to its usability. Some scale studies in the literature (Stes et al., 2010) provide evidence of differences between cultures. However, in some studies conducted in different cultures, it has also been observed that scales developed in some cultures produce similar results in other cultures (Schellhase, 2009; Tezci, 2017; Zhang, 2001). Demes and Geeraert (2014) determined in their research conducted in nine different cultures that they supported a similar factor structure and obtained reliable results. Beaton et al. (2000) stated that different results could be obtained in scale studies conducted in culturally different countries due to translation-related issues and depending on temporal changes.

In curriculum assessment studies, teachers' opinions regarding their participation in in-service training have been analysed. The analysis results indicate that the averages in all factors are significantly higher for teachers who participated in in-service training than those who did not. Teachers who receive education related to the curriculum are expected to evaluate it better, and their knowledge about implementing it is expected to be higher. Erdogan et al. (2005) emphasized in their study the importance of in-service training in teachers' curriculum implementation and curriculum-related assessments. Erss (2018) pointed out the differences in curriculum assessment in a study conducted with teachers from three countries. Sjögrén et al. (2003) stated that teacher experiences impact curriculum evaluation. Thompson et al. (2013) also emphasized the importance of knowledge and experience in teachers' curriculum implementation and evaluation. The difference in teachers' in-service training levels regarding curriculum-related evaluations is noteworthy in indicating the extent to which they know about this curriculum.

6. Conclusion

The scale developed in this study is based on teachers' evaluations in a specific culture's curriculum studies. However, the scale developed in this study was prepared regardless of cultural context. Therefore, it was prepared considering the characteristics of a curriculum implementation, structure, and function in general, not specific to the curriculum context in Kosovo. In this context, the scale in question will likely increase its applicability in different cultures. The scale, which is based on teachers' self-reports, was developed differently than the scales available in the literature. The items have been written considering elements such as content, learning outcomes, curriculum philosophy, assessment and assessment as one dimension and features such as understandability, coherence, relevance to students, functionality as the qualities that elements such as content, learning outcomes planning etc. should include assessment and evaluation represent another dimension.

The scale is compatible in terms of its factors and the items within the factors. The scale is compatible in terms of its factors and the items within the factors. In this framework, the scale can be used both in the training of teachers and prospective teachers and in curriculum development studies and in the evaluation of the curriculum in practice. and program design preparation.

7. Suggestions

Conducting further research to test the applicability and reliability of the scale in different cultural contexts will provide to determine whether the scale remains valid in different educational settings and take into account any cultural nuances that may influence the assessment process. Increasing

the sample size and diversity of participants in future studies will also strengthen the generalizability of the results. Including a wider range of teachers with different levels of education and backgrounds will provide a more comprehensive understanding of the effectiveness of the scale. Finally, conducting comparative studies between different curriculum and regions using the scale is recommended in future research. Such studies can provide insights into the strengths and weaknesses of various curriculum and serve as a guide for improvements and best practices in curriculum development and implementation.

Author contributions: Each author made an equal contribution to the current study and has read and given their approval to the article's final published version.

Declaration of interest: The authors declared that there were no potential conflicts of interest.

Data availability: The datasets generated during and/or analysed during the current study are available from the corresponding author on reasonable request.

Ethical statement: All participants voluntarily provided informed consent for their involvement in the study, including the collection and analysis of data. Given the nature of the study and the procedures followed, no further ethical approval was required.

Funding: No funding source is reported for this study.

References

Adams, J. E. (2000). *Taking charge of curriculum: Teacher networks and curriculum implementation.* Teachers College Press.

Akıncı, M., & Köse, E. (2021). Research trends of program evaluation studies conducted between 2010-2019 in Turkey. *Cukurova University Faculty of Education Journal, 50*(1), 77-120.

Apsari, Y. (2018). Teachers' problems and solutions in implementing curriculum 2013. *Acuity: Journal of English Language Pedagogy, Literature and Culture, 3*(1), 11-23. https://doi.org/10.35974/acuity.v3i1.620

Bartlett, M. S. (1951). The effect of standardization on a Chi-square approximation in factor analysis. *Biometrika, 38*(3/4), 337-344. https://doi.org/10.1093/biomet/38.3-4.337

Beaton, D. E., Bombardier, C., Guillemin, F., & Ferraz, M. B. (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

Ben-Chaim, D., Joffe, N., & Zoller, U. (1994). Empowerment of elementary school teachers to implement science curriculum reforms. *School Science and Mathematics, 94*(7), 356-366. https://doi.org/10.1111/j.1949-8594.1994.tb15694.x

Bentler, P. M., & Bonett, D. G. (1980). Significance tests and goodness of fit in the analysis of covariance structures. *Psychological Bulletin, 88*, 588-606. https://doi.org/10.1037/0033-2909.88.3.588

Bledsoe, K. L., & Graham, J. A. (2005). The use of multiple evaluation approaches in program evaluation. *American Journal of Evaluation, 26*(3), 302-319. https://doi.org/10.1177/1098214005278749

Briggs, N.E., & MacCallum. R.C. (2003). Recovery of weak common factors by maximum likelihood and ordinary least squares estimation. *Multivariate Behavioral Research, 38*(1). 25-56. https://doi.org/10.1207/S15327906MBR3801_2

Brown, J. D. (1995). *The elements of language curriculum - A systematic approach to program development.* Heinle & Heinle Publishers.

Browne, M. W. (2001). An overview of analytic rotation in exploratory factor analysis. *Multivariate Behavioral Research, 36*(1), 111-150. https://doi.org/10.1207/S15327906MBR3601_05

Burul, C., & Tezci, E. (2022). A scale development study to determine teachers' curriculum fidelity. *Journal of National Education, 51*(235), 2417-2446. https://doi.org/10.37669/milliegitim.896628

Button, L. J. (2021). *Curriculum essentials: A journey*. Pressbooks.

Byrne, B. M. (2001). *Structural equation modeling with AMOS*. Lawrence Erlbaum Associates.

Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin, 56*(2), 81-105. https://doi.org/10.1037/h0046016

Carless, D. R. (1998). A case study of curriculum implementation in Hong Kong. *System, 26*(3), 353-368. https://doi.org/10.1016/S0346-251X(98)00023-2

Cattell, R. B. (1978). *The scientific use of factor analysis in behavioral and life sciences.* Plenum. https://doi.org/10.1007/978-1-4684-2262-7

Clark, C., &. Yinger, R. J. (1987). Teacher planning. In D. Berliner, & B. Rosenshine (Eds.), *Talks to teachers* (pp. 342-365). Random House.

Colton, D., & Covert, R. (2007). *Designing and constructing ınstruments for social research and evaluation.* Jossey-Bass.

Comrey, A.L. & Lee, H. B. (1992). *A first course in factor analysis.* Erlbaum.

Connelly, F. M. (1980). Teachers' roles in the using and doing of research and curriculum development. *Journal of Curriculum Studies, 12*(2), 95-107. https://doi.org/10.1080/0022027800120202

Dagenais, M. E., Hawley, D., & Lund, J. P. (2003). Assessing the effectiveness of a new curriculum: Part I. *Journal of dental education, 67*(1), 47-54. https://doi.org/10.1002/j.0022-0337.2003.67.1.tb03618.x

DASH. (2001). *Korniza e kurrikulit të ri të kosovës* [The framework of the new curriculum of Kosovo]. Author.

Demes, K. A., & Geeraert, N. (2014). Measures matter: Scales for adaptation, cultural distance, and acculturation orientation revisited. *Journal of Cross-Cultural Psychology, 45*(1), 91-109. https://doi.org/10.1177/0022022113487590

Dindar, H., & Yaygın, S. (2007). Teachers' perceptions about the transition process to elementary school science and technology teaching curriculum. *Kastamonu Education Journal, 15*(1), 185-198.

Ding, L., Velicer, W. F. & Harlow, L. L. (1995). Effects of estimation methods, number indicators per factor, and improper solutions on structural equation modeling fit indices. *Structural Equation Modeling, 2,* 119-144. https://doi.org/10.1080/10705519509540000

Dunn, T. J., Baguley, T., & Brunsden, V. (2014). From alpha to omega: A practical solution to the pervasive problem of internal consistency estimation. *British Journal of Psychology, 105*(3), 399-412. https://doi.org/10.1111/bjop.12046

Ediger, M. (2003). *Philosophy and curriculum.* Discovery Publishing House.

Elliott, J. (1994). The teacher's role in curriculum development: An unresolved issue in English attempts at curriculum reform. *Curriculum Studies, 2*(1), 43-69. https://doi.org/10.1080/0965975940020103

Erdoğan, M., Kayır, Ç. G., Kaplan, H., Ünal, Ü. Ö. A., & Akbunar, Ş. (2005). Teachers views on curriculum developed since 2005: A content analysis of the researches between 2005 and 2011. *Kastamonu Education Journal, 23*(1), 171-196.

Erss, M. (2018). 'Complete freedom to choose within limits'–teachers' views of curricular autonomy, agency and control in Estonia, Finland and Germany. *The Curriculum Journal, 29*(2), 238-256. https://doi.org/10.1080/09585176.2018.1445514

Evans, B. (2003). *The William and Mary curriculum user survey results*. Center for Gifted Education.

Evans, W. (1986). An investigation of curriculum implementation factors. *Education, 106*(4), 447-453.

Fabrigar, L. R., Wegener, D. T., MacCallum, R. C., & Strahan, E. J. (1999). Evaluating the use of exploratory factor analysis in psychological research. *Psychological Methods, 4*(3), 272–299. https://doi.org/10.1037/1082-989X.4.3.272

Ford, J. K., McCallum, R. S. & Tait, M. (1986). The application of exploratory factor analysis in applied psychology: A critical review and analysis. *Personnel Psychology, 39,* 291314. https://doi.org/10.1111/j.1744-6570.1986.tb00583.x

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

Fu, Y., & Sibert, S. (2017). Teachers' perspectives: Factors that impact implementation of integrated curriculum in K-3 classrooms. *International Journal of Instruction, 10*(1), 169-186. https://doi.org/10.12973/iji.2017.10111a

Gomez, R., & Fisher, J. W. (2003). Domains of spiritual well-being and development and validation of the Spiritual Well-Being Questionnaire. *Personality and Individual Differences, 35*(8), 1975–1991. https://doi.org/10.1016/S0191-8869(03)00045-X.

Goodlad, J. I. (1964). School curriculum reform in the United States. In D. Flinders & S. Thornton (Eds.), *The Curriculum studies reader* (pp.45-54). Routledge Falmer.

Gouveia, V. V. & Soares, A. K. S. (2015). *Calculadoras de validade de construto* [Construct validity calculators]. Universidade Federal da Paraíba.

Gredler, M. E. (1996). *Program evaluation.* Prentice Hall.

Gutek, G. L. (1988). *Philosophical and ideological voices on education.* Prentice-Hall.
Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2019). *Multivariate data analysis.* Cengage Learning.
Harrison, D.A. & McLaughlin, M.E. (1993). Cognitive processes in self-report responses: Tests of item context effects in work attitude measures. *Journal of Applied Psychology, 78,*129-140. https://doi.org/10.1037/0021-9010.78.1.129
Heling, Z., & Bangxiu, X. (2009). The philosophical foundation and practice of the reform in the contemporary curriculum and instruction. *Tattva Journal of Philosophy, 1*(2), 57-71. https://doi.org/10.12726/tjp.2.5
Heyman, R. D. (1981). Analyzing the curriculum. *International Review of Education, 27,* 449-470. https://doi.org/10.1007/BF00598141
Ho, R. (2006). *Handbook of univariate and multivariate data analysis and interpretation with SPSS.* Chapman & Hall. https://doi.org/10.1201/9781420011111
Horn, J. L. (1965). A rationale and test for the number of factors in factor analysis. *Psychometrika, 30*(2), 179-185. https://doi.org/10.1007/BF02289447
Hu, L. T., & Bentler, P. M. (1999). Cut-off criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling, 6,* 1-55. https://doi.org/10.1080/10705519909540118
Kaiser, H. F. (1974). An index of factorial simplicity. *Psychometrika, 39*(1), 31-36. https://doi.org/10.1007/BF02291575
Kandemir, M. A., Tezci, E., Shelley, M., & Demirli, C. (2019). Measurement of creative teaching in mathematics class. *Creativity Research Journal, 31*(3), 1-12. https://doi.org/10.1080/10400419.2019.1641677
Karakus, G. (2021). A literary review on curriculum implementation problems. *Shanlax International Journal of Education, 9*(3), 201-220. https://doi.org/10.34293/education.v9i3.3983
Kelly, A. L. (2004). *The curriculum- theory and practice.* Sage.
Kern, D. E., Thomas, P. A., & Hughes, M. T. (2007). *Curriculum development for medical education: a six-step approach.* The John's Hopkins University Press.
Kline, P. (1994). *An easy guide to factor analysis.* Routledge.
Ladesma, R.D. & Valero- Mora, P. (2007). Determining the number of factors to retain in EFA: An easy-to-use computer program for carrying out parallel analysis. *Practical Assessment, Research and Evaluation, 12*(2), 1-11.
Lewthwaite, B. (2001). *The development, validation and application of a primary school science curriculum implementation questionnaire* [Unpublished Doctoral dissertation]. Curtin University, Perth.
MacCallum, R. C., Browne, M. W., & Cai, L. (2007). Factor analysis models as approximations. In R. Cudeck & R. C. MacCallum (Eds.). *Factor analysis at 100: Historical developments and future directions* (pp.153-175). Lawrence Erlbaum.
MacCallum, R. C., Widaman, K. F., Zhang, S., & Hong, S. (1999). Sample size in factor analysis. *Psychological Methods, 4*(1), 84–99. https://doi.org/10.1037/1082-989x.4.1.84
Macdonald, D. (2003). Curriculum change and the post-modern world: Is the school curriculum reform movement an anachronism? *Journal of Curriculum Studies, 35*(2), 139-149. https://doi.org/10.1080/00220270210157605
Male, B. (2012). *The primary curriculum design handbook: preparing our children for the 21st century.* Continuum International Publishing.
Malhotra, N. K. (2011). *Pesquisa de marketing: Uma orientação aplicada* [Marketing research: An applied orientation]. Bookman.
Maren, M. S., Salleh, U. K. M., & Zulnaidi, H. (2021). Assessing prospective teachers' soft skills curriculum implementation: Effects on teaching practicum success. *South African Journal of Education, 41*(3), Article 1915. https://doi.org/10.15700/saje.v41n3a1915
Marsh, C. J. & Willis, G. (2007). *Curriculum: Alternative approaches, ongoing issues.* Merrill Prentice Hall.
MAShT. (2016a). *Korniza e kurrikules e arsimit parauniversitar të Republikës së Kosovës* [Pre-university education curriculum framework of the Republic of Kosovo]. Blendi.
MAShT. (2016b). *Kurrikula bërthamë për arsimin e mesëm të ulët-Klasa VI,VII,VIII,IX* [Core curriculum for lower secondary education-Grades VI, VII, VIII, IX]. Blendi.
McCormick, R., & James, M. (2018). *Curriculum evaluation in schools.* Routledge. https://doi.org/10.4324/9780429454233
Mehmeti, F., & Tezci, E. (2018). Evaluation of the sixth grade technology teaching curriculum in Kosovo. *Turkish Studies, 13*(11), 933-960.

Nevenglosky, E. A. (2018). *Barriers to effective curriculum implementation* [Unpublished Doctoral dissertation]. Walden University, Minneapolis.

Nouraey, P., Al-Badi, A., Riasati, M. J., & Maata, R. L. (2020). Educational program and curriculum evaluation models: A mini systematic review of the recent trends. *Universal Journal of Educational Research, 8*(9), 4048-4055. https://doi.org/10.13189/ujer.2020.080930

Nunnally, J.C. & Bernstein, I. H. (1994). *Psychometric theory.* McGraw-Hill.

O'Connor, B. P. (2000). SPSS and SAS programs for determining the number of components using parallel analysis and Velicer's MAP test. *Behavior Research Methods, Instrumentation, and Computers, 32,* 396-402.

Ornstein, A. C. (1990). Philosophy as a basis for curriculum decisions. *The High School Journal, 74*(2), 102-109.

Ornstein, A. C., & Hunkins, F. P. (2003). *Kurrikula - baza, parime dhe probleme* [Curriculum, foundations, principles and issues]. Institute of Pedagogical Studies.

Patton, M. Q. (2002). *Qualitative research and evaluation methods.* Sage.

Posner, G. J., & Strike, K. A. (1974). *An analysis of curriculum structure* (ED089432). ERIC.

Pratt, D. (1994). *Curriculum planning.* Hardcourt Brace College.

Raykov, T. (1997). Estimation of composite reliability for congeneric measures. *Applied Psychological Measurement, 21*(2), 173-184. https://doi.org/10.1177/01466216970212006

Şahan, H. H. (2007). *Evaluation of third year mathematics program in elementary school* (Publication no. 229042) [Doctoral dissertation, Hacettepe University]. Council of Higher Education Thesis Center.

Saylor, J. & M. Alexander, W. (1973). *Planning curriculum for schools.* Holt, Rinehert and Winston.

Schellhase, K. C. (2009). *Are approaches to teaching and/or student evaluation of instruction scores related to the amount of faculty formal educational course work* [Unpublished doctoral dissertation]. University of Central Florida Orlando, Florida.

Scott, D. (2001). *Curriculum and assessment.* Greenwood Publishing.

Sjögrén, A., Poskiparta, M., Liimatainen, L., & Kettunen, T. (2003). Teachers' views on curriculum development in health promotion in two Finnish polytechnics. *Nurse Education Today, 23*(2), 112-122. https://doi.org/10.1016/S0260-6917(02)00165-X

Stes, A., De Maeyer, S., & Van Petegem, P. (2010). Approaches to teaching in higher education: Validation of a Dutch version of the Approaches to Teaching Inventory. *Learning Environments Research, 13*(1), 59-73. http://dx.doi.org/10.1007/s10984-009-9066-7

Stufflebeam, D. (2001). Evaluation models. *New Directions for Evaluation, 89,* 7-98. https://doi.org/10.1002/ev.3

Taba, H. (1962). *Curriculum development: Theory and practice.* Harcourt, Brace and World.

Tabachnick, B. G., & Fidel, L. S. (2001). *Using multivariate statistics.* Allyn and Bacon.

Tanner, D., &. Tanner, L. N. (1970). *Curriculum development.* Macmillan.

Tezci, E. (2017). Adaptation of ATI-R Scale to Turkish samples: Validity and reliability analyses. *International Education Studies, 10*(1), 67-81. https://doi.org/10.5539/ies.v10n1p67

Thompson, B. (2004). *Exploratory and confirmatory factor analysis: Understanding concepts and applications.* American Psychological Association. https://doi.org/10.1037/10694-000

Thompson, D., Bell, T., Andreae, P., & Robins, A. (2013). The role of teachers in implementing curriculum changes. In T. Camp & P. Tymann (Eds.), *Proceeding of the 44th ACM Technical Symposium on Computer Science education* (pp. 245-250). SIGCSE. https://doi.org/10.1145/2445196.2445272

Van den Akker J. (2003). Curriulum perspetives: An introdution. In J. van den Akker, W. Kuiper & U. Hameyer (Eds.), *Curriculum landscapes and trends* (pp. 1-10). Kluwer. https://doi.org/10.1007/978-94-017-1205-7_1

Van den Akker, J., Gravemeijer, K., McKenney, S., & Nieveen, N. (2006). *Educational design research.* Routledge. https://doi.org/10.4324/9780203088364

VanTassel-Baska, J., & Brown, E. F. (2007). Toward best practice: An analysis of the efficacy of curriculum models in gifted education. *Gifted Child Quarterly, 51*(4), 342-358. https://doi.org/10.1177/0016986207306323

Velicer, W. F. (1976). Determining the number of components from the matrix of partial correlations. *Psychometrika, 41,* 321-327. https://doi.org/10.1007/BF02293557

Velicer, W. F., Eaton, C. A., & Fava, J. L. (2000). Construct explication through factor or component analysis: A review and evaluation of alternative procedures for determining the number of factors or components. In R. D. Goffin & E. Helmes (Eds.), *Problems and solutions in human assessment: Honoring Douglas N. Jackson at seventy* (pp. 41–71). Kluwer Academic/Plenum Publishers. https://doi.org/10.1007/978-1-4615-4397-8_3

Walker, F. D. (2003). *Fundamentals of curriculum passion and professionalism.* Lawrence Erlbaum Associates. https://doi.org/10.4324/9781410606259

Wang, H., & Cheng, L. (2009). Factors affecting teachers' curriculum implementation. *Linguistics Journal, 4*(2), 135-166.

Williams, B., Onsman, A., & Brown, T. (2020). Exploratory factor analysis: A five-step guide for novices. *Journal of Emergency Primary Health Care, 8*(3), 1-13. https://doi.org/10.33151/ajp.8.3.93

Yavuz, G., & Doğan, N. (2015). Using Velicer's Map Test and Horn's Parallel Analysis for determining component number. *Hacettepe University Journal of Education, 30*(3), 176-188.

Zhang, L. F. (2001). Approaches and thinking styles in teaching. *The Journal of Psychology, 135*(5), 547-561. http://dx.doi.org/10.1080/00223980109603718.

Zwick, W. R., & Velicer, W. F. (1986). Comparison of five rules for determining the number of components to retain. *Psychological Bulletin, 99*(3), 432–442. https://doi.org/10.1037/0033-2909.99.3.432