Necatibey Faculty of Education Electronic Journal of Science and Mathematics Education
Vol. 18, Issue 2, December 2024, pp. 503-533. ISSN: 1307-6086

Research Article

A Scale Development Study to Use the Flipped Classroom Model in Mathematics Education*

Emine ÖZDEMİR[1], Beyzanur TOPÇU GÖRE[2]

[1] Balıkesir University, Necatibey Faculty of Education, Türkiye, eozdemir@balikesir.edu.tr, http://orcid.org/0000-0002-4114-0005

[2] Balıkesir Karesi District Directorate of National Education, Türkiye, beyzanet26@gmail.com, http://orcid.org/0000-0003-2826-5601

Received : 29.05.2024 Accepted : 27.11.2024

Doi: https://doi.org/10.17522/balikesirnef.1488745

Abstract – In this study, it was aimed to develop a valid and reliable scale to evaluate the use of the Flipped Classroom Model in secondary school mathematics teaching. The study design was determined as a survey design. In the process of developing the scale, an item pool was first created and expert opinion was obtained. Adjustments were made according to the opinions of experts. The study group was determined by simple random sampling and the application was carried out with a total of 871 middle school students from grades 5, 6, 7, and 8. According to the data obtained, validity and reliability analyses of the scale were conducted with SPSS and AMOS programs. The Flipped Classroom Model Evaluation (FCME) Scale consists of two subscales: the "Evaluation of Teaching Conducted with Video Lessons" (ETCVL) subscale and the "Evaluation of Teaching Conducted with Face-to-Face Lessons" (ETCFFL) subscale. As a result of the analyses, the FCME scale consisting of 32 items in total was developed. The reliability coefficient of the ETCVL subscale was .846, the reliability coefficient of the ETCFFL subscale was .775 and the reliability coefficient of the scale as a whole was .886. As a result of the analysis, it was determined that the FCME Scale is valid and reliable.

Keywords: Flipped classroom model, math education, secondary school students, scale development.

Corresponding author: Emine ÖZDEMİR , eozdemir@balikesir.edu.tr

*This study includes a part of the second author's Master's thesis entitled "A Scale Development Study to Use the Flipped Classroom Model" conducted under the supervision of the first author.

Introduction

With the rapid development of science and technology, it is known that the use of smartphones, tablets, computers, etc. has increased and products such as videos, photographs, and designs are created using these tools and shared on various social platforms (Kocaman Karoğlu, 2015). In societies exposed to this situation, it is known that there is an increase in curiosity about technology, active use of technology, and an increase in the desire to improve their knowledge and skills on how to use technology. While technology is used so actively in daily life, the fact that educational environments lag in the use of technology has led to digital incompatibility (Atal & Koçak Usluel, 2011).

The quality of the education provided is directly proportional to the quality of the education provided by teachers blending technology, pedagogy, and field knowledge. In the literature, the blending of these three knowledge is called Technological Pedagogical Content Knowledge (TPACK) (Mutluoğlu & Erdoğan, 2016). When we look at the common intersection set of the components, TPACK defines the course designs prepared by integrating technological-pedagogical-field knowledge. (Koehler & Mishra, 2009). In the updated curricula by the Ministry of National Education (MoNE), the necessity of imparting eight key competencies to students during the educational processes is mentioned, including digital competence, competence in science and technology, learning to learn, and mathematical competence (MoNE, 2018).

The FC Model is a learning model that emerged to present the subject to the student with formats such as slides, videos, interactive content, etc. on digital platforms so that students who cannot come to class for various reasons do not fall behind (Baker, 2000). According to this learning model, students learn the subject through digital platform content and then come to the educational environment to reinforce the subject with activities. At first, this method was used only for students who could not come to class, but after receiving positive feedback, the researchers started to apply this method to the whole class. With this method, while saving the time allocated for lecturing in the lesson, it also created an opportunity for students who lacked knowledge or missed the subject to complete the subject (Bender, 2018; Bergmann, 2011).

When the studies in the literature are examined, it is seen that the lessons taught with the FC model increase students' academic achievement, active participation in the classroom, self-regulation skills, retention of the subject, teacher-student communication, make the lesson more understandable, fun and remarkable, and students begin to develop positive attitudes towards mathematics lessons (Arslan, 2021; Bolatlı, 2018; Bulut, 2019; Güç, 2017; Houston, 2020; Özdemir, 2016; Wei et al., 2020).

When the studies on the use of the FC Model were examined, it was found that there was no study in which students from all grade levels were studied at the secondary school level, and that interview forms were used as data collection tools to determine the opinions of students, and that there was no scale with proven validity and reliability. It is thought that the research conducted in this direction can contribute to these deficiencies identified in the literature.

It was determined that there were few studies on the use of the FC Model in secondary school mathematics courses and the problem situations (attitudes, opinions, academic achievement, etc.) examined in the studies were similar. Considering the study group, it was observed that the study was generally conducted with a single grade level. Upon examination, it was determined that there was no scale for the evaluation of the FC Model and that there was no study conducted with all grade levels at the same time. Accordingly, this study aims to develop a scale with proven validity and reliability for middle school students to evaluate the use of the FC model in mathematics education and to contribute to the deficiencies in the literature.

Method

Research Design

As the aim of this study is to develop a scale, the research design is a survey design from quantitative research methods. The survey design enables the quantitative presentation of the attitudes, opinions or tendencies of the population by conducting studies with a sample selected from the population (Creswell, 2017) and obtaining information by reaching many participants (Büyüköztürk et al., 2018).

Participants

The schools and classes to be included in the study group were determined by a simple random sampling method. According to this sampling method, each individual is equally likely to be selected for sampling and it is the most powerful and valid method compared to other sampling methods in terms of representing the universe (Büyüköztürk et al., 2018).

The study group of the research included a total of 871 secondary school students from six public schools in Sancaktepe district of Istanbul province, studying in grades 5-6-7-8 in the second semester of the 2020-2021 academic year. Of these 871 students, 512 were female and 359 were male. In terms of grade level, there are 236 students in 5th grade, 289 students in 6th grade, 200 students in 7th grade, and 146 students in 8th grade.

Data Collection

Process of Scale Development

In the process of scale development in line with the purpose of the study, the process in Figure 1 was followed based on Yurgudül's (2005) scale development scheme.

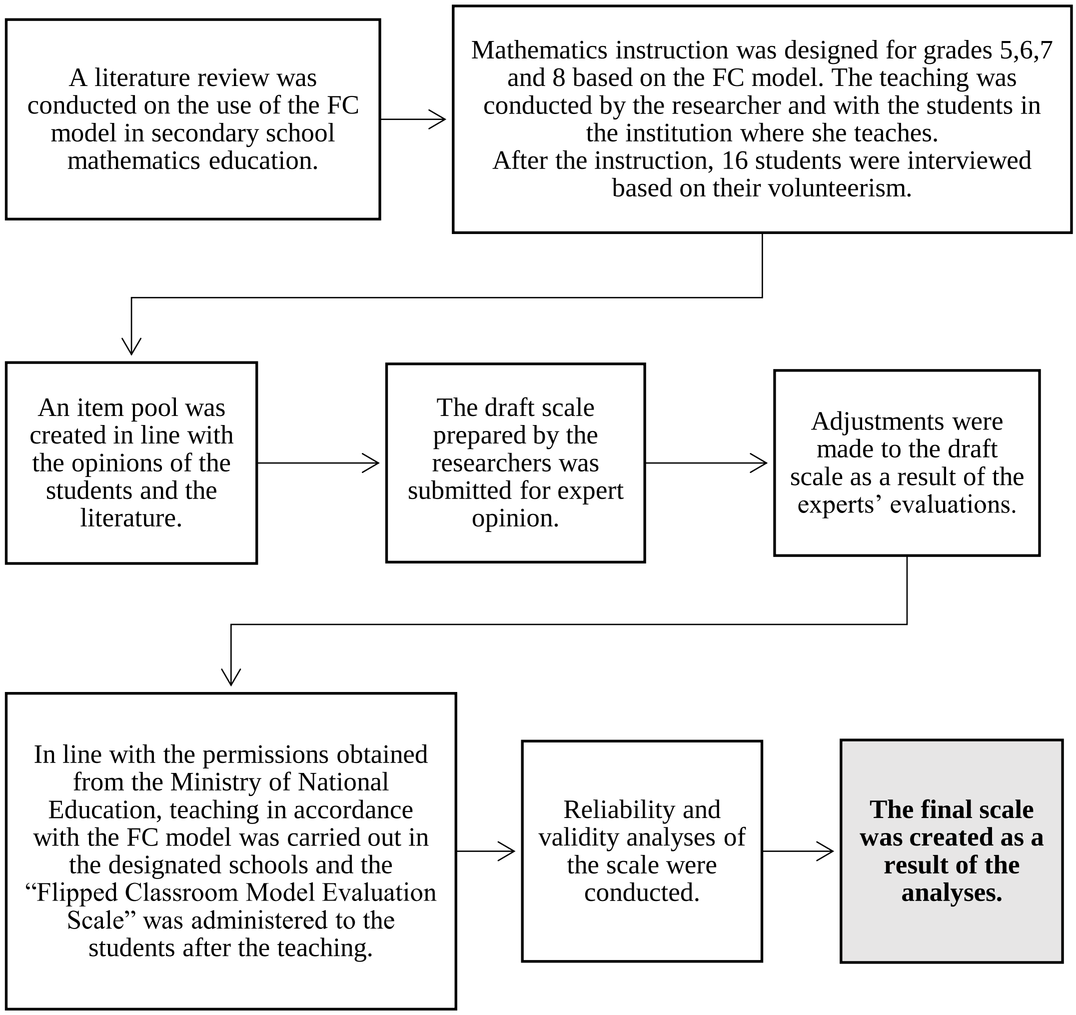

Figure 1 Scale Development Process

In the national literature, graduate thesis studies on scale development were examined through the YÖKTEZ database. Since this study is a scale development study for the evaluation of a FC model and the AMOS program of structural equation modeling was used, the combinations of the keywords “flipped, scale development, structural equation, AMOS” were searched by keeping the keyword “scale development” constant.

The literature review was limited with the criteria of being between 2012 and 2022 to cover the last 10 years and being in the field of education and training. When the words “flipped, scale development, AMOS”, “flipped, scale development, structural equation” were searched in the thesis name, abstract, subject and index sections of the theses from the detailed search section provided by YÖKTEZ, no study was found. Then, when the words “flipped, scale development” were searched, three thesis studies in which a scale was developed related to the FC model were found and one article was found as a result of the literature review. When the words “structural equation, scale development” were searched, 14 thesis studies were found. Six of these studies were excluded from the literature review because they did not use structural equation modeling during scale development.

In addition to these studies, the study titled “Scale development process in the field of education: a content analysis study” by Şahin and Boztunç Öztürk (2018), which examined 69 articles developed in the field of education in Türkiye between 2010 and 2016, was also included in the review. For this reason, a total of 13 studies were analyzed under the title of “studies on scale development”.

In the literature review, four scales related to the flipped classroom model were found. One of these studies is Durak's (2017) adaptation of the scale developed by Hao (2016) to measure students' readiness into Turkish. In his study, Kurtoğlu (2019) used the scale adapted into Turkish by Durak (2017) and adapted it to teachers by conducting reliability studies. Erensayın (2019) developed self-efficacy and perception scales for teachers regarding the applicability of the FC model; Akgün (2015) developed a scale aiming to examine the effect of the FC model on students' academic achievement and opinions.

When the thesis studies in which the structural equation model was used among the scale development studies in the field of education were examined, it was determined that the analysis program used varied. It can be said that the use of different analysis programs such as AMOS and LISREL in the studies depends on the preference of the researchers.

In the thesis study conducted by Balkaya (2022), it was aimed to develop the “YouTube Usage Scale” for middle school students. For this purpose, a total of 644 students from 5th, 6th, 7th and 8th grades were included in the study. The draft scale, which was initially prepared with 47 items, was reduced to 42 items after expert opinion. Then, EFA and CFA analyses were conducted with SPSS and AMOS programs. As a result of these analyzes, the final scale was determined as 25 items. The reliability coefficient of the final scale was found to be .91.

In the study conducted by Nanto (2021), the effect of error management culture and job attraction behaviors on organizational creativity was examined. Accordingly, the researcher developed the “Error Management Culture” scale for quantitative data. The sample of the study consisted of 747 primary and secondary school teachers. For the scale, a 41-item draft scale was first created as a result of the literature review. The number of items was reduced to 26 by removing 15 items from the draft scale sent to expert opinion. Then, EFA and CFA analyses conducted with SPSS and AMOS programs resulted in a final scale consisting of 16 items. The reliability coefficient of the scale was found to be .846.

The validity and reliability analyses of the scales developed by Erensayın (2019), Kurtoğlu (2019) and Durak (2017) were included in the studies. However, in the scale study developed by Akgün (2015) for the evaluation of the FC model, it was found that no information on validity and reliability analyses was provided.

The draft scale was prepared in a 5-point Likert scale and named the “Flipped Classroom Model Evaluation (FCME) Scale”. Due to the structure of the FC model, teaching takes place in two phases: “video lesson teaching” and “face-to-face teaching”. Due to this structure of the FC model, the FCME scale consists of two subscales, namely the “Evaluation of Teaching Conducted with Video Lessons (ETCVL) Subscale” and the “Evaluation of Teaching Conducted with Face-to-Face Lessons (ETCFFL) Subscale”. This nomenclature was chosen because the scale consists of two subscales.

Process of Creating the Substance Pool

In accordance with the decision taken at the meeting of Balıkesir University Science and Engineering Sciences Ethics Committee dated 20.04.2021 and numbered 2021/2, it was approved to collect the data of the study.

In the training given to create an item pool, content was prepared in Edpuzzle according to the gains below and teaching was carried out.

- M.5.3.1.2. Collects data related to research questions and shows them in a frequency table and column graph.

- M.6.1.5.3. Makes multiplication of a natural number and a fraction and makes sense of it.
- M.6.1.5.4. Makes multiplication of two fractions and makes sense of it.
- M.7.1.3.2 Multiplies and divides rational numbers.
- M.8.1.3.4. Performs multiplication and division operations with square roots.

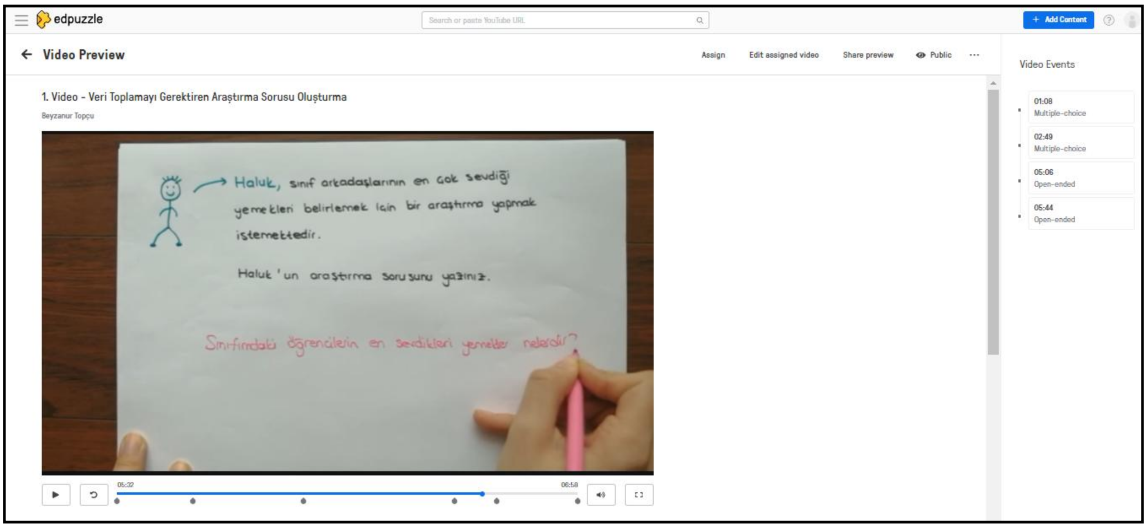

Figure 2 An Example of Video Lessons

After the instruction, the step of creating an item pool for the scale was started. At this stage, volunteer students were contacted to get their opinions on the evaluation of teaching with the FC model. The following open-ended questions were asked to the 16 students who volunteered to give their opinions and the opinions of the students were obtained.

- "Has studying the topic beforehand had any effect on the desire to participate in solving problems during the lesson? If so, what kind of impact did it have?
- What are your thoughts on interactive questions in video lessons?
- What are your thoughts on feedback provided for interactive questions in video lessons?
- What do you think about using Edpuzzle and online quizzes/games at the end of video lessons? What are the positive and negative aspects?
- How did you feel using a site like Edpuzzle, Wordwall, Quizizz, etc., which you had never used before?

- Can you make a comparison between math lessons taught with the FC model and those not taught with it?"

"In this application, I think it is better because there is not a specific time like a lecture because some of us may be in another place, for example, they may not have internet, so they cannot open it, so it is good that there is not a specific time, they watch it the next day and then they go to class (Student C)"

Based on this student's opinion, items 1, 2, and 21 of the ETCVL subscale were created.

"When someone gets it right, you write the correct answer bravo, and when they get it wrong, they can see why they got it wrong. I think this is very nice (Student N)."

Similarly, Student N's opinions helped to write the 4th and 8th items of the ETCVL Subscale.

When examining the qualitative data resulting from the interviews, it was observed that students responded to the "evaluation of teaching conducted with video lessons", "evaluation of teaching conducted face-to-face", "the use of technology in mathematics instruction", and "their attitudes towards mathematics" within the FC model. When the responses from the students were analyzed, the idea that the evaluation of teaching should consist of two stages emerged.

These two stages are named "Evaluation of Teaching Conducted with Video Lessons" and "Evaluation of Teaching Conducted Face-to-Face," and accordingly, an item pool has been created. At the stage of creating the item pool, the scale was prepared in 5-point Likert type from multiple scales. In the Likert scale, various statements are directed to the sample and it is expected to determine the degree of agreement or disagreement with these statements (Altunışık et al., 2005). After reviewing the responses, the draft scale was prepared as 90 items. These items were prepared and organized using the Microsoft Form tool.

Content Validity

It was aimed to ensure the content validity of the scale by applying expert opinion. A form was prepared in which "Appropriate, Should be improved, Not appropriate" options were added to each of the scale items. If the option "should be improved" was selected, a text box was opened and the experts were asked to indicate their suggestions.

The draft scale was sent via e-mail to a total of 11 experts, including 3 mathematics and 1 Turkish teacher with at least 5 years of experience, 1 measurement and evaluation, 1 Turkish education, 3 mathematics education, and 2 educational experts from other fields.

The content validity index of the draft scale was calculated based on the opinions of the experts. The Content Validity Ratio (CVR) approach is used in scale development studies to quantify the opinions of experts regarding the content validity of the items in the developed scale (Yurdugül & Bayrak, 2012). For each item in the draft scale, a calculation was made using the content validity ratio formula.

Since the number of experts was 11, the content validity criterion for each item was determined as .59 (Yurdugül, 2005). The CVR value was found to be higher than .59 for each item on the scale. With the content validity index (CVI), a value is calculated for the entire scale. CVI is calculated by averaging the CVR values of the items decided to be included in the scale (Yurdugül & Bayrak, 2012). The experts evaluated the 90 items from different perspectives, such as the FC model, grammar, suitability for the purpose, comprehensibility of the items, and the presence of similar items. Edits were made by taking into account the opinions of the experts and 26 of the 90 items were deleted and a draft scale consisting of 64 items was formed. Accordingly, at $\alpha = .05$ level of significance, the content validity index value of the FCME scale was calculated as 0.98 and it was concluded that the content validity of the scale had a high value.

Scoring for the scale items was done as 1=Strongly disagree, 2=Disagree, 3=Neither agree nor disagree, 4=Agree, 5=Strongly agree for the positive items, and the reverse for the negative items.

Students who scored high on the scale can be said to have positive evaluations of mathematics teaching according to the FC model.

Teaching Process Based on the FC Model and Collection of Scale Data

In the six public secondary schools included in the study, mathematics teaching based on the FC model was implemented. Before the implementation, care was taken to ensure that there were students from all levels (5th, 6th, 7th, and 8th grades) of each secondary school included in the implementation. Instruction was carried out with a total of 871 students and then the instruction was evaluated by applying the draft scale.

In the process of teaching based on the FC model, content was prepared according to the following acquisitions and teaching was carried out.

- M.5.2.2.3.Determines and draws the elements of rectangle, parallelogram, rhombus and trapezoid.
- M.6.3.2.2.Form the area relation of a parallelogram and solve related problems.
- M.7.3.2.1.Explain the side and angle properties of regular polygons.
- M.7.3.2.2.Determines the diagonals, interior and exterior angles of polygons, calculates the sum of the measures of interior and exterior angles.
- M.8.3.1.3.Relates the side lengths of a triangle to the measures of the angles opposite these sides.

Data Analysis

Structural Equation Modeling was preferred in the data analysis of the study. Structural Equation Modeling (SEM) adopts a confirmatory approach by analyzing all direct or indirect effects between variables at the same time, while also taking into account the errors of the variables (Demir, 2016).

After the implementation of mathematics teaching according to the FC model, the draft scale was applied to the students. The data received after the application were transferred to the computer environment. Data analysis was conducted using the SPSS 24.0 program for factor analysis, t-test, and reliability analysis, and the AMOS 24.0 program for factor analysis within the scope of SEM.

Validity analyses (exploratory factor analysis, confirmatory factor analysis) and reliability analyses (item analysis, t-test, Cronbach's alpha coefficient) were conducted separately for the two subscales of the FCME scale, namely, "Evaluation of Teaching Conducted with Video Lessons (ETCVL) subscale" and "Evaluation of Teaching Conducted with Face-to-Face Lessons (ETCFFL) Subscale".

Results

Analyses Related to the Evaluation of Teaching Conducted with Video Lessons (ETCVL) Subscale

The ETCVL subscale was determined as 49 items after expert opinion. First, the Kaiser Meyer Olkin (KMO) value was calculated to determine the suitability of this subscale for factor analysis. Since the KMO value of the ETCVL subscale was found to be .93, it was determined that the scale data were suitable for factor analysis.

Five exploratory factor analyses were applied to the ETCVL subscale and a total of 21 items with loadings below .50 or overlapping items were removed from the scale.

Table 1 Results of the Rotated Components Matrix of the Items in the ETCVL Subscale

	Rotated component matrix					
	Factors					
Items	1	2	3	4	5	6
videolesson14	.720	.021	-.054	.106	.076	-.016
videolesson13	.708	.028	.054	.046	.033	-.054
videolesson12	.706	.031	.156	.069	-.010	-.037
videolesson23	.682	.060	.011	.083	-.039	-.122
videolesson26	.676	.112	.291	.122	-.003	-.014
videolesson25	.659	.191	.191	.168	.057	-.029
videolesson22	.652	.083	.215	.127	.018	-.030
videolesson16	.626	-.119	-.007	.209	.149	.061
videolesson11	.612	.016	.152	.178	.085	.021
videolesson9	.579	.005	.141	.092	.267	.096
videolesson10	.573	.059	.188	.092	.183	.071
videolesson4	.558	.083	.090	.039	.252	-.101
videolesson18	.262	.723	-.196	-.022	-.216	.007
videolesson30	.096	.713	-.010	.057	-.037	.120
videolesson19	.223	.643	-.019	.059	-.174	.114
videolesson35	.153	.643	.136	-.125	-.050	.107
videolesson29	-.155	.610	.222	-.067	.188	.077
videolesson31	-.279	.594	.138	.027	.233	.040
videolesson34	.253	.065	.792	.096	.026	-.017
videolesson37	.193	.081	.781	.161	-.031	-.033
videolesson33	.407	.055	.587	.084	.048	.029
videolesson40	.176	-.069	.127	.750	.022	.066
videolesson38	.224	.032	.221	.685	.070	-.147
videolesson39	.258	-.011	.000	.668	.162	-.002
videolesson2	.213	-.022	.071	.070	.746	-.053
videolesson3	.265	-.085	-.082	.162	.677	-.010
videolesson45	-.009	.198	-.056	-.022	.067	.799
videolesson47	-.090	.179	.032	-.024	-.127	.787

It was determined that the lowest loading values of the items constituting the scale were .558 and the highest .799.

The items belonging to the ETCVL subscale were coded as “videolesson...”. In addition, it was found that the subscale of the ETCVL consisted of six factors. These factors were named “positive attitude” (items 4, 9, 10, 11, 12, 13, 14, 16, 22, 23, 25, 26), “negative attitude” (items 18, 19, 29, 30, 31, 35), “self-efficacy” (items 33, 34 and 37), "advantages" (items 38, 39 and 40), "accessibility" (items 2 and 3) and "safety" (items 45 and 47).

Confirmatory factor analysis was conducted to determine the construct validity of the remaining 28 items in the ETCVL subscale. As a result of the analysis, items 4, 29 and 31 were removed from the scale because their standardized regression weights were below .5. Improvement was made by creating covariance between items with high error percentages. Table 2 shows the fit values obtained as a result of these procedures.

Table 2 Confirmatory Factor Analysis Fit Values of the Items in the ETCVL Subscale

Model fit indices		Good fit	Acceptable fit	Found value
CMIN/DF		x^2/ sd ≤ 3	x^2/ sd ≤ 5	2.520
Comparative fit indices	NFI	.95 ≤ NFI	.90 ≤ NFI	.906
	CFI	.97 ≤ CFI	.95 ≤ CFI	.941
	RMSEA	RMSEA ≤ .05	RMSEA ≤ .08	.042
Absolute fit indices	GFI	.90 ≤ GFI	.85 ≤ GFI	.942
Residual-Based Fit Indices	RMR	0 < RMR ≤ .05	0 < RMR ≤ .08	.041

According to the fit values of x^2/df = 2.520 < 3, .90 ≤ NFI = .906, .95 ≤ CFI = .941, RMSEA = .042 ≤ .05, .90 ≤ GFI = .942, RMR = .041 ≤ .05 for the final scale items of the ETCVL subscale, there is an excellent fit between the model and the data (Karagöz, 2021).

Based on the findings obtained as a result of the analyses, the validity of the final six-factor version of the ETCVL subscale was proved.

The AMOS diagram of the ETCVL subscale obtained as a result of confirmatory factor analysis is shown in Figure 2.

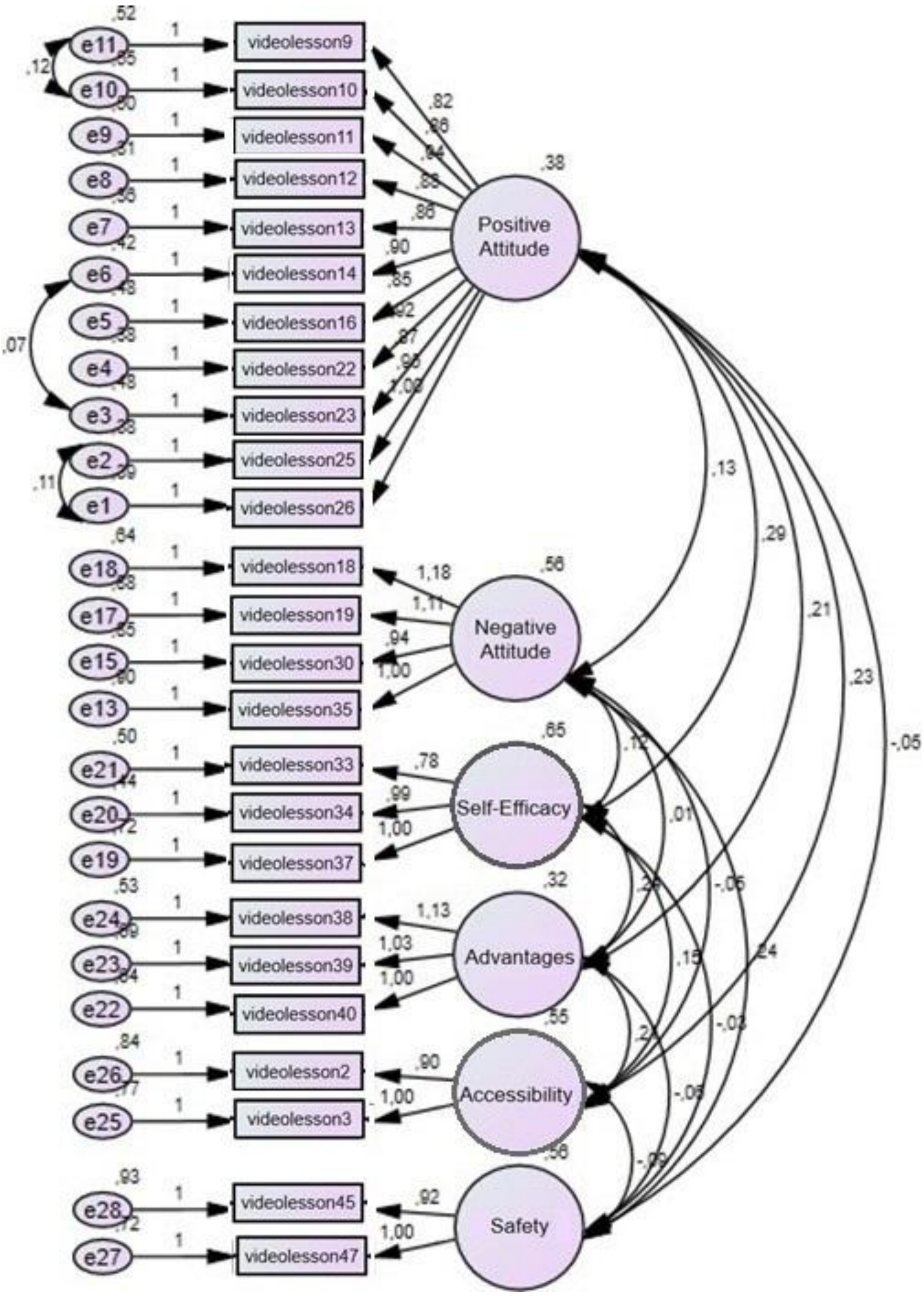

Figure 2 AMOS Diagram of the Items in the ETCVL Subscale

Analyses Related to the Evaluation of Teaching Conducted with Face-to-Face Lessons (ETCFFL) Subscale

The ETCFFL subscale was determined as 15 items after expert opinion. First, the Kaiser Meyer Olkin (KMO) value was calculated to determine the suitability of this subscale for factor analysis. Since the KMO value of the ETCFFL subscale was found to be .851, it was determined that the scale data were suitable for factor analysis.

Three exploratory factor analyses were applied to the ETCFFL subscale and a total of 6 items with loadings below .45 or overlapping items were removed from the scale.

Table 3 Results of the Rotated Components Matrix of the Items in the ETCFFL Subscale

	Rotated component matrix	
	Factors	
Items	1	2
facetofacelesson7	.781	.231
facetofacelesson6	.759	.168
facetofacelesson5	.722	.229
facetofacelesson10	.704	.057
facetofacelesson13	.600	-.168
facetofacelesson1	.548	.394
facetofacelesson11	.545	-.109
facetofacelesson2	.025	.891
facetofacelesson3	.058	.882

It was determined that the lowest loading values of the items constituting the scale were .545 and the highest .891.

The items belonging to the ETCFFL subscale were coded as “facetofacelesson...”. In addition, it was found that the subscale of the ETCFFL consisted of two factors. These factors were named “positive attitude” (items 1, 5, 6, 7, 10, 11, and 13) and “negative attitude” (items 2 and 3).

Confirmatory factor analysis was conducted to determine the construct validity of the remaining 9 items in the ETCFFL subscale. As a result of the analysis, items 11 and 13 were removed from the scale because their standardized regression weights were below .5. Improvement was made by creating covariance between items with high error percentages. Table 2 shows the fit values obtained as a result of these procedures.

Table 4 Confirmatory Factor Analysis Fit Values of the Items in the ETCFFL Subscale

Model fit indices		Good fit	Acceptable fit	Found value
CMIN/DF		x^2/ sd ≤ 3	x^2/ sd ≤ 5	4.117
	NFI	.95 ≤ NFI	.90 ≤ NFI	.976
Comparative fit indices	CFI	.97 ≤ CFI	.95 ≤ CFI	.982
	RMSEA	RMSEA ≤ .05	RMSEA ≤ .08	.060
Absolute fit indices	GFI	.90 ≤ GFI	.85 ≤ GFI	.984
Residual-based fit indices	RMR	0 < RMR ≤ .05	0 < RMR ≤ .08	.039

According to the fit values of x^2/df = 4.117 < 5, .95 ≤ NFI = .976, .95 ≤ CFI = .982, .05 ≤ RMSEA = .060, .90 ≤ GFI = .984, RMR = .039 ≤ .05 for the final scale items of the ETCFFL subscale, there is an excellent fit between the model and the data (Karagöz, 2020).

Based on the findings obtained as a result of the analyses, the validity of the final two-factor version of the ETCFFL subscale was proved.

The AMOS diagram of the ETCFFL subscale obtained as a result of confirmatory factor analysis is shown in Figure 3.

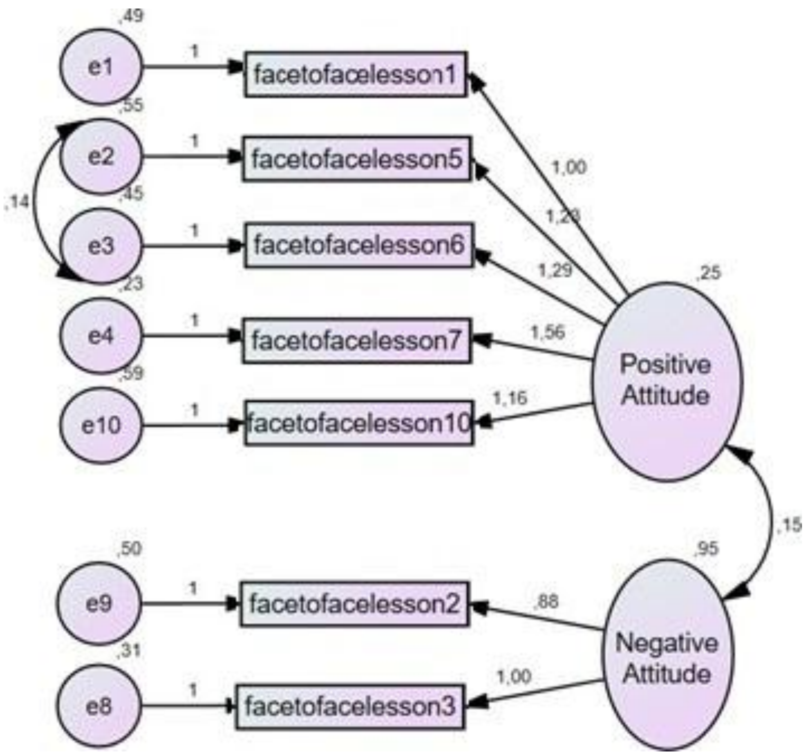

Figure 3 AMOS Diagram of the Items in the ETCFFL Subscale

As a result of the exploratory and confirmatory factor analyses of the FCME scale,

- Twenty-four of the 49 items in the ETCVL subscale were removed from the scale, leaving 25 items in the final scale. Items 2, 3, 9, 10, 11, 12, 13, 14, 16, 18, 19, 22, 23, 25, 26, 30, 33, 34, 35, 37, 38, 39, 40, 45, 47 remained in the final version of ETCVL subscale.
- Eight of the 15 items in the ETCFFL subscale were removed from the scale and 7 items remained in the final scale. Items 1, 2, 3, 5, 6, 7 and 10 remained in the final version of the ETCFFL subscale

Table 5 and Table 6 show the factors formed in the two subscales of the final version of the FCME scale and the items included in these factors.

Table 5 The Factors and Items in the ETCVL Subscale and the Items in These Factors

Evaluation of teaching conducted with video lessons subscale	
Factor 1: Positive attitude	
videolesson9	Even if I mark the wrong answer in interactive questions, I can easily learn the correct solution thanks to the feedback.
videolesson10	I like the feedback in the interactive questions to find out where I went wrong.
videolesson11	There must be an interactive question in the video lesson.
videolesson12	I wonder where I went wrong in interactive questions.
videolesson13	Seeing where I went wrong in interactive questions makes it easier for me to understand the topic.
videolesson14	I like to see an encouraging statement that my answer to the interactive questions is correct.
videolesson16	Feedback is necessary in interactive questions.
videolesson22	When I do the activities at the end of the video lesson, I also realize which subjects I lack knowledge about.
videolesson23	I would like to see my name at the top of the points ranking of the activities at the end of the video lesson.
videolesson25	The activities at the end of the video lesson help me to reinforce the topic.
videolesson26	The activities at the end of the video lesson increase the retention of my learning.
Factor 2: Negative attitude	
videolesson18	When I cannot solve a question in interactive questions, I stop watching the video lesson.
videolesson19	Learning the subject from video lessons is boring.
videolesson30	I do not continue the activity at the end of the video lesson when I cannot meet the time limit.
videolesson35	I only watch video lessons because my teacher gives me homework.
Factor 3: Self-efficacy	
videolesson33	I always search for the correct solutions to the questions I cannot solve in the activities at the end of the video lesson.
videolesson34	I watched the video lesson several times before coming to the face-to-face lesson.
videolesson37	I can watch the video lessons over and over again without getting bored.
Factor 4: Advantages	
videolesson38	It allows me to determine the most appropriate time for my learning environment.
videolesson39	It increases my level of active participation in the lesson.
videolesson40	It allows me to manage my own learning process.
Factor 5: Accessibility	
videolesson2	I can watch video lessons wherever I want (at home, outside, on the go, etc.).
videolesson3	I can watch video lessons whenever I want.
Factor 6: Safety	
videolesson45	I do not find it safe to use unfamiliar learning platforms.
videolesson47	I do not want to provide my personal information to learning platforms, even for educational purposes.

Table 6 The Factors and Items in the ETCFFL Subscale and the Items in These Factors

Evaluation of teaching conducted with face-to-face lessons subscale	
Factor 1: Positive Attitude	
facetofacelesson1	At the beginning of a face-to-face lesson, my teacher summarizes the topic, which helps me to remember what I have learned.
facetofacelesson5	I learn better when I attend a face-to-face lesson by watching a video lesson.
facetofacelesson6	I feel more confident when I attend a face-to-face lesson by watching a video lesson.
facetofacelesson7	I would like to participate more in face-to-face lessons as I have knowledge about the subject by watching videos.
facetofacelesson10	I would like to be the first person to answer questions when I attend a face-to-face class by watching video lessons.
Factor 2: Negative Attitude	
facetofacelesson2	It is a waste of time for my teacher to summarize what I have learned in video lessons in face-to-face lessons.
facetofacelesson3	I get bored when my teacher summarizes the topic in video lessons at the beginning of a face-to-face lesson.

When the items in the final version of the FCME scale that emerged as a result of the explanatory and confirmatory factor analyses were examined, it was found that 24 items were positive and 8 items were negative. Items 10, 11, 16, 19, 24 and 25 in the ETCVL subscale and items 2 and 3 in the ETCFFL subscale are negative and the remaining items are positive (Appendix 1).

Reliability Analyses

Item analysis, the differences between the item mean scores of the lower 27% and upper 27% groups formed according to the total scores of the scale were examined using an unrelated t-test. If the differences observed between the groups are significant, it can be interpreted that the internal consistency of the scale is ensured. According to the results of item analysis, it shows to what extent the items distinguish individuals in terms of the measured behavior (Büyüköztürk, 2019).

The FCME scale consists of two different scales, namely the evaluation of teaching conducted with video lessons (ETCVL) subscale and the evaluation of teaching conducted with face-to-face lessons (ETCFFL) subscale. Reliability analyses were therefore conducted separately for both subscales.

With the analyses performed here, a) t-test for the significance of the differences between the item mean scores of the lower 27% and upper 27% groups to be formed according to the scale total scores, b) the reliability of the scale items using item-total correlations, and c) the reliability of the scale using Cronbach's Alpha.

Table 7 Item Analysis Results of the Items in the Subscales of the ETCVL and ETCFFL

Item no	Item-total correlation[1]	t (lower %27 – upper %27)[2]
videolesson2	.246	-8.010
videolesson3	.235	-7.290
videolesson9	.497	-.511
videolesson10	.519	-15.227
videolesson11	.535	-15.633
videolesson12	.572	-16.301
videolesson13	.537	-15.158
videolesson14	.546	-16.678
videolesson16	.488	-13.481
videolesson18	.323	-13.725
videolesson19	.368	-15.832
videolesson22	.580	-16.418
videolesson23	.538	-16.647
videolesson25	.646	-19.164
videolesson26	.638	-19.356
videolesson30	.312	-12.564
videolesson35	.344	-14.850
videolesson37	.409	-13.868
videolesson34	.458	-14.692
videolesson40	.325	-9.304
videolesson38	.391	-11.386
videolesson39	.352	-10.168
videolesson45	.064	-4.322
videolesson47	.021	-3.515
videolesson33	.516	-16.777
facetofacelesson1	.518	-15.376
facetofacelesson2	.322	-14.064
facetofacelesson3	.391	-15.492
facetofacelesson5	.558	-17.650
facetofacelesson6	.596	-18.707
facetofacelesson7	.624	-19.656
facetofacelesson10	.496	-15.630

[1] $n = 871$,
[2] $n_1 = n_2 = 236$,
***$p < .001$

Table 7 shows that the item-total correlations for all items in the ETCVL subscale ranged between .021 and .646 and the t-values were significant ($p<.001$), while the item-total correlations for all items in the ETCFFL subscale ranged between .322 and .624 and the t-values were significant ($p<.001$). These results suggest that the reliability of the items in the

ETCVL and ETCFFL subscales is high and that they are discriminative and intended to measure the same behavior.

Table 8 shows the Cronbach’s alpha and McDonald's Omega reliability coefficients of the subscales of the ETCVL and ETCFFL.

Table 8 Cronbach’s Alpha and McDonald's Omega Values of the FCME Scale and Subscales

FCME scale	Cronbach’s Alpha reliability coefficient	McDonald's Omega reliability coefficient
ETCVL subscale	.846	.840
ETCFFL subscale	.775	.735
Total	.886	.879

Regarding the reliability of the FCME scale, Cronbach’s alpha coefficient was found to be .846 for the ETCVL subscale, .775 for the ETCFFL subscale, and .886 for the total scale. Since these values are greater than .70 (Büyüköztürk, 2019), it was found that the FCME scale is reliable (Table 8). In addition, since McDonald's Omega (ω) values are above .70, it is understood that the reliability of the scale is sufficient and structural reliability is also provided (McDonald, 1985).

Table 9 shows Cronbach’s alpha and McDonald's Omega reliability coefficients of the factors belonging to the ETCVL subscale and the ETCFFL subscale of the FCME scale.

Table 9 Cronbach’s Alpha and McDonald's Omega Values of the Subscales and Factors of the Subscales

Subscales	Factors	Cronbach’s Alpha reliability coefficient	McDonald's Omega reliability coefficient
Evaluation of teaching conducted with video lessons (ETCVL)	Positive attitude	.851	.888
	Negative attitude	.751	.752
	Self-efficacy	.740	.752
	Advantage	.630	.630
	Accessibility*	.548	-
	Safety*	.555	-
Evaluation of teaching conducted with face-to-face lessons (ETCFFL)	Positive attitude	.813	.816
	Negative attitude*	.804	-

* Omega cannot be estimated because the number of items is less than 3.

The Cronbach’s alpha reliability coefficients of the positive attitude, negative attitude, and self-efficacy factors of the ETCVL subscale were found to be .851, .751, and .740, respectively. In addition, McDonald's Omega (ω) values for the factors of the subscales ranged between .630 and .888.

Since these values were between .70 and .90, these three factors were accepted as reliable. The Cronbach's alpha reliability coefficient of the advantage factor was found to be .630, which is acceptable as it is between .60 and .70. The Cronbach's alpha reliability coefficients of the accessibility and safety factors were calculated as .548 and .555, respectively. Since the reliability coefficients of the reliability factor are between .50 and .60, it is known that the reliability of this factor is weak. However, since the items in this factor are important and necessary for the evaluation of the teaching according to the FC model, it was deemed appropriate to keep it as it is.

The Cronbach's alpha reliability coefficients of the positive attitude and negative attitude factors of the ETCFFL subscale were calculated as .813 and .804, respectively. Since these values were between .70 and .90, these two factors were accepted as reliable.

Conclusion, Discussion, and Suggestions

This study aims to develop a scale for middle school students' evaluation of mathematics teaching according to the FC model. According to the FC model, students are taught the subject with a video lesson before coming to class, and during class, the subject is summarized and reinforced with various activities. Therefore, the FCME scale developed within the scope of the research consists of two subscales, namely the "Evaluation of Teaching Conducted with Video Lessons" (ETCVL) subscale and the "Evaluation of Teaching Conducted with Face-to-Face Lessons" (ETCFFL) subscale

In the first stage of the scale development process, a literature review was conducted. Within the scope of the information obtained, mathematics teaching was designed according to the FC model, and a sample application was made to middle school students. The opinions of the students about the model were obtained by interviewing the volunteer students who took part in the implementation. An item pool consisting of 90 items was obtained in line with the opinions of the students. This item pool was sent to 11 experts and expert opinions were received. Edits were made by taking expert opinions into consideration and a draft scale consisting of 64 items was obtained. A pilot study was conducted with this draft scale. Validity and reliability analyses were conducted according to the data obtained after the application and a final 5-point Likert-type scale consisting of 32 items in total was obtained.

Şahin and Boztunç Öztürk (2018) examined the scale development studies according to the "group size" variable and found that only 12 articles out of 69 articles had group sizes in the range of 500-999. Based on this result, it can be said that this study, in which 871 students participated, is one of the few scale development studies.

Şahin and Boztunç Öztürk (2018) examined scale development studies according to the "ratio of participants per item". The number of articles with a ratio of 20 and above is only 4 out of 69 articles. In this study, the ratio of participants per item was 27.22 (871:32), making it one of the few scale development studies in the literature. Similarly, the number of articles using "exploratory factor analysis, confirmatory factor analysis" among the methods of determining construct validity is 32 out of 69 articles. 19 articles use sub-upper group analysis and item-test correlation in item analysis and 45 articles examine internal consistency in determining reliability.

In this scale development study, internal consistency was examined to determine reliability, and Cronbach alpha values were calculated. According to the results of EFA and CFA, the ETCVL subscale consists of 25 items and a 6-factor structure. The factors in the ETCVL subscale are named positive attitude, negative attitude, self-efficacy, advantage, accessibility, and safety when analyzed in terms of the items they contain. The reliability coefficient of the ETCVL subscale was calculated as .846 and it was concluded that it was reliable. When the reliability coefficients of the factors were calculated,

- The positive attitude factor consists of 11 items and its reliability coefficient was found to be .851,
- The negative attitude factor consists of 4 items and its reliability coefficient was found to be .751,
- The self-efficacy factor consists of 3 items and its reliability coefficient was found to be .740,
- The advantage factor consists of 3 items and its reliability coefficient was found to be .630,
- The accessibility factor consists of 2 items and its reliability coefficient was found to be .548 and
- The safety factor consists of 2 items and the reliability coefficient was found to be .555.

The reliability coefficients of the advantage, accessibility, and safety factors were lower than .70. However, it was decided to keep these items and factors in the scale because the items in these factors were considered to be necessary for the scale, they could not be gathered under a single factor in terms of the characteristics they measured, and the final version of the ETCVL subscale was calculated to be reliable.

According to the results of EFA and CFA, the ETCFFL subscale consists of 7 items and a 2-factor structure. When the common characteristics of the items that make up the factors in the ETCFFL subscale were examined, these factors were named positive attitude and negative attitude.

The reliability coefficient of the ETCFFL subscale was calculated as .775 and it was concluded that it was reliable. When the reliability coefficients of the factors are calculated;

- The positive attitude factor consists of 5 items and its reliability coefficient was found to be .813, and
- The negative attitude factor consists of 2 items and the reliability coefficient was found to be .804.

After calculating the reliability coefficients of the two subscales of the FCME scale, the reliability coefficient of the scale as a whole was found to be .886 and it was concluded that the FCME scale was reliable.

There are 24 positive and 8 negative items in the FCME scale. The scores that can be obtained from the scale vary between 32 and 160.

In the literature review, four scales related to the FC model were found. Of these scales, two are scale development studies (Akgün, 2015; Erensayın, 2019), one is the adaptation of an existing scale in the literature into Turkish (Durak, 2017), and the other is the adaptation of a scale for students in the literature for teachers (Kurtoğlu, 2019). When these four studies were evaluated in terms of disciplines, it was determined that they were not in the field of mathematics. Within this scope, no scale was found in the literature for the evaluation of the use of the FC model in mathematics teaching. In conclusion, a valid and reliable scale for the evaluation of the use of the FC model in mathematics teaching was developed in this study.

The validity and reliability analyses of the scales developed by Erensayın (2019), Kurtoğlu (2019), and Durak (2017) are available. Nevertheless, it was found that the scale developed by Akgün (2015) did not provide information on validity and reliability analyses. Detailed information on the reliability and validity analyses of the FCME scale was given and the scale was found to be reliable and valid.

Akgün (2015) developed a scale titled “Opinion Questionnaire on flipped classrooms”. However, it is important that the study did not include validity and reliability studies. Furthermore, the scale items evaluated the “inverted-flat classroom practice” in general, which negatively affects the content validity of the scale. In this study, unlike Akgün (2015), there are more detailed items for the evaluation of teaching and items that allow the evaluation of every dimension of teaching. Another difference is that in this study, the two phases of the teaching of the FC model (teaching with video lessons and teaching with face-to-face lessons) were handled separately. Studies also differ in terms of the online platforms used in the teaching phase with video lessons.

The scale items developed in Erensayın’s (2019) study and the scale items developed in this study have similar items. The difference between this study and Erensayın’s (2019) study is that there is a negative attitude sub-dimension belonging to both stages of the instruction.

It was determined that the study conducted by Durak (2017) and this study were similar in terms of items. The difference between this study and Durak’s (2017) study is that there are negative items in the scale developed and it is a scale for the evaluation of two important stages of teaching.

When the scale development studies were evaluated in terms of the study group, they were conducted with secondary school teachers from different branches (Erensayın, 2019), middle school teachers from different branches, and middle school (5th and 8th grade) students (Kurtoğlu, 2019), middle school 5th and 6th-grade students (Durak, 2017), and middle school 5th-grade students (Akgün, 2015). This study was conducted with all grade levels (5th, 6th, 7th, and 8th grade students). In this way, a scale was created for all grade levels at the secondary school level.

Erensayın (2019) used exploratory factor analysis in the scale development study and Durak (2017) used confirmatory and exploratory factor analysis in the study. In the study conducted by Durak (2017), although it is not specified which program was used, it is seen that the visuals and fit criteria of the Structural Equation Model were used for the confirmatory factor. In this direction, it can be said that the scale of LCME, whose validity was increased with the AMOS statistical program compatible with the Structural Equation Model, was introduced to the field.

In conclusion of this study, suggestions for researchers are given below.

- This scale can be used to evaluate the teaching of mathematics according to the FC model.
- The scale can also be used for primary school, high school, and university students by conducting a reliability study.
- Since the items in the scale are general statements, they can be used for the evaluation of instruction for each learning area or sub-learning area of mathematics.
- Although the FCME scale was prepared for the mathematics course, it does not contain mathematical terms, so it can be used in the evaluation of teaching based on the FC model for different branches.
- This scale can be used in the studies to be conducted by teachers, lecturers, and prospective teachers on education.

Compliance with Ethical Standards

Disclosure of potential conflicts of interest

There was no conflict of interest in this study.

Funding

No financial support was received.

CRediT author statement

This study is a part of the master's thesis of the second author under the supervision of the first author.

Research involving Human Participants and/or Animals

This research was conducted with the permission obtained in accordance with the decision taken at the meeting of the Science and Engineering Sciences Ethics Committee dated 20.04.2021 and numbered 2021/2

Matematik Eğitiminde Ters Yüz Edilmiş Sınıf Modelinin Kullanımına Yönelik Bir Ölçek Geliştirme Çalışması

Özet:

Bu araştırmada ortaokul matematik öğretiminde Ters Yüz Edilmiş Sınıf Modelinin kullanımının değerlendirilmesini sağlayacak geçerli ve güvenilir bir ölçek geliştirmek amaçlanmıştır. Araştırmanın deseni tarama deseni olarak belirlenmiştir. Ölçeğin geliştirilme sürecinde ilk olarak madde havuzu oluşturulmuş ve uzman görüşü alınmıştır. Uzmanların görüşlerine göre düzenlemeler yapılmıştır. Çalışma grubu basit seçkisiz örneklemeyle belirlenmiş olup 5,6,7 ve 8.sınıflardan olmak üzere toplam 871 ortaokul öğrenciyle uygulama gerçekleştirilmiştir. Elde edilen verilere göre ölçeğin geçerlik ve güvenirlik analizleri SPSS ve AMOS programlarıyla yapılmıştır. Ters Yüz Edilmiş Sınıf Modeli Değerlendirme(TYESMD) ölçeği, "Video Ders İle Yapılan Öğretimin Değerlendirilmesi(VDYÖD) Alt Ölçeği" ve "Yüz Yüze Ders İle Yapılan Öğretimin Değerlendirilmesi(YDYÖD) Alt Ölçeği" olmak üzere iki alt ölçekten oluşmaktadır. Yapılan analizler sonucunda toplamda 32 maddeden oluşan TYESMD ölçeği geliştirilmiştir. VDYÖD alt ölçeğinin güvenirlik katsayısı .846, YDYÖD alt ölçeğinin .775 ve bir bütün olarak ölçeğin güvenirlik katsayısı .886 olarak bulunmuştur. Yapılan analizler sonucunda TEYSMD Ölçeğinin geçerli ve güvenilir bir ölçek olduğu tespit edilmiştir.

Anahtar kelimeler: Ters yüz edilmiş sınıf modeli, matematik eğitimi, ortaokul öğrencileri, ölçek geliştirme.

References

Akgün, M. (2015). *The effect of flipped classroom on students' academic achievement and views* (Publication No. 423423) [Master's thesis, Fırat University]. Council of Higher Education National Thesis Center.

Altunışık, R., Coşkun, R., Bayraktaroğlu, S., & Yıldırım, E. (2005). *Sosyal bilimlerde araştırma yöntemleri*: Spss *uygulamalı* (4th ed.). Sakarya.

Arslan, U. (2021). *Investigation of the effect the flipped classroom model on middle school students' academic achievement and self-regulation skills in mathematics course* (Publication No. 695917) [Master's thesis, Çukurova University]. Council of Higher Education National Thesis Center.

Atal, D., & Koçak Usluel, Y. (2011). Elementary school students' use of technology in and out of school. *Hacettepe University Journal of Education*, 41, 24-35. https://dergipark.org.tr/en/download/article-file/87383

Baker, J. W. (2000). The "classroom flip": Using web course management tools to become the guide by the side. In: Chambers JA (ed.), *Selected Papers from the 11th International Conference on College Teaching and Learning (*pp. 9–17*)*. Florida Community College at Jacksonville. https://upcea.edu/wp-content/uploads/2020/09/The-Classroom-Flip-Baker.pdf

Balkaya, M. (2022). *Developing and implementing a measurement tool related to Youtube use with learning purposes for middle school student* (Publication No. 731719) [Master's thesis, Çanakkale Onsekiz Mart University]. Council of Higher Education National Thesis Center.

Bender, W.N. (2018). Ters yüz edilmiş stem sınıfı. In S. Durmuş, A. S. İpek, & B. Yıldız (Trans. ed.), *STEM öğretimi için 20 strateji* (1st ed., pp. 73-82). Nobel.

Bergmann, J. (2011, May 6). The history of the flipped class. *Flippedclass.com* https://flippedclass.com/the-history-of-the-flipped-class/

Bolatlı, Z. (2018). *Determining the academic achievement of students who use flipped classroom method supported by mobile application and their views on cooperative learning* (Publication No. 534559) [Master's thesis, Selçuk University]. Council of Higher Education National Thesis Center.

Bulut, R. (2019). *Investigation of the use of the flipped clasroom model in teaching the ratio-proportion* (Publication No. 563939) [Master's thesis, Erzincan Binali Yıldırım University]. Council of Higher Education National Thesis Center.

Büyüköztürk, Ş. (2019). *Sosyal bilimler için veri analizi el kitabı istatistik, araştırma deseni spss uygulamaları ve yorum* (25th ed.). Pegem.

Büyüköztürk, Ş., Kılıç Çakmak, E., Akgün, Ö.E., Karadeniz Ş., & Demirel, F. (2018). *Bilimsel araştırma yöntemleri* (24th ed.). Pegem.

Creswell, J. W. (2017). *Araştırma deseni nitel, nicel ve karma yöntem yaklaşımları* (3rd ed.). (S.B., Demir, Trans.). Eğiten Kitap.

Demir, B. (2016). *A scale development for anxiety of student for finding work with structural equation model: An application faculty of economics and administrative sciences in Cumhuriyet University* (Publication No. 459470) [Master's thesis, Cumhuriyet University]. Council of Higher Education National Thesis Center.

Durak, H. Y. (2017). Turkish adaptation of the flipped learning readiness scale for middle school students. *Bartın University Journal of Faculty of Education, (6)*3, 1056-1068. https://doi.org/10.14686/buefad.328826

Erensayın, E. (2019). *Examination of the applicability of the flipped classroom model in secondary education according to teacher perceptions* (Publication No. 580481) [Master's thesis, Van Yüzüncü Yıl University]. Council of Higher Education National Thesis Center.

Güç, F. (2017). *The effect of the flipped classroom practice on the rational numbers and operations with rational numbers* (Publication No. 478696) [Master's thesis, Amasya University]. Council of Higher Education National Thesis Center.

Hao, Y. (2016). Middle school students' flipped learning readiness in foreign language classrooms: Exploring its relationship with personal characteristics and individual circumstances. *Computers in Human Behavior, 59,* 295-303. https://doi.org/10.1016/j.chb.2016.01.031

Houston, K. (2020). *The flipped classroom's impact on math anxiety, self-efficacy, and motivation at the middle school level* (Publication No. 28260171) [Doctoral dissertation, Cabrini University]. ProQuest Dissertations & Theses Global.

Karagöz, Y. (2021). *SPSS AMOS META uygulamalı biyoistatistik* (3rd ed.). Nobel.

Koehler, M., & Mishra, P. (2009). What is technological pedagogical content knowledge (Tpack)?. *Contemporary Issues in Technology and Teacher Education, 9*(1), 60-70. https://www.learntechlib.org/primary/p/29544/

Kocaman Karoğlu, A. (2015). The changing nature of storytelling by means of technology in the instructional process: digital storytelling. *Educational Technology Theory and Practice*, *5*(2), 89-106. https://doi.org/10.17943/etku.29277

Kurtoğlu, C. (2019). *Investigation of the readiness of middle school teachers and students regarding the flipped classroom model* (Publication No. 605297) [Master's thesis, Fırat University]. Council of Higher Education National Thesis Center.

McDonald, R. P. (1985). *Factor analysis and related methods* (1st ed.). Erlbaum. https://doi.org/10.4324/9781315802510

MoNE (2018). *Matematik dersi öğretim programı (ilkokul ve ortaokul 1, 2, 3, 4, 5, 6, 7 ve 8. sınıflar)*. http://mufredat.meb.gov.tr/ProgramDetay.aspx?PID=329

Mutluoğlu, A., & Erdoğan, A. (2016). Examining primary mathematics teachers' technological pedagogical content knowledge (TPACK) levels according to their preferred teaching styles. *OPUS International Journal of Society Researches*, *6*(10), 102-126. https://dergipark.org.tr/en/pub/opus/issue/22701/242324#article_cite

Nanto, Z. (2021). *The relationship between error management culture, work engagement and organizational creativity* (Publication No. 659550) [Doctoral dissertation, Fırat University]. Council of Higher Education National Thesis Center.

Özdemir, A. (2016). *Flipped classroom model practice focused on blended learning in secondary mathematics education* (Publication No. 429471) [Doctoral dissertation, Fırat University]. Council of Higher Education National Thesis Center.

Şahin, M. G., & Boztunç Öztürk, N. (2018). Scale Development Process in Educational Field: A Content Analysis Research. *Kastamonu Education Journal*, 26(1), 191-199. https://doi.org/10.24106/kefdergi.375863

Wei, X., Cheng, I., Chen, N. S., Yang, X., Liu, Y., Dong, Y., & Zhai, X. (2020). Effect of the flipped classroom on the mathematics performance of middle school students. *Educational Technology Research and Development*, *68*, 1461-1484. https://doi.org/10.1007/s11423-020-09752-x

Yurdugül, H. (2005, September 28-30). *Ölçek geliştirme çalışmalarında kapsam geçerliği için kapsam geçerlik indekslerinin kullanılması*. XIV. Ulusal Eğitim Bilimleri Kongresi, Pamukkale Üniversitesi Eğitim Fakültesi, Denizli. http://yunus.hacettepe.edu.tr/~yurdugul/3/indir/PamukkaleBildiri.pdf

Yurdugül, H., & Bayrak, F. (2012). Content validity measures in scale development studies: Comparison of content validity index and kappa statics. *Hacettepe University Journal of Education*, Special Issue 2, 264-271. http://efdergi.hacettepe.edu.tr/yonetim/icerik/makaleler/655-published.pdf

Appendix 1

The Flipped Classroom Model Evaluation (FCME) Scale

	Evaluation of Teaching Conducted with Video Lessons (ETCVL) Subscale	Strongly Disagree	Disagree	Neither agree nor disagree	Agree	Strongly Agree
1	I can watch video lessons wherever I want (at home, outside, on the go, etc.).					
2	I can watch video lessons whenever I want.					
3	Even if I mark the wrong answer in interactive questions, I can easily learn the correct solution thanks to the feedback.					
4	I like the feedback in the interactive questions to find out where I went wrong.					
5	There must be an interactive question in the video lesson.					
6	I wonder where I went wrong in interactive questions.					
7	Seeing where I went wrong in interactive questions makes it easier for me to understand the topic.					
8	I like to see an encouraging statement that my answer to the interactive questions is correct.					
9	Feedback is necessary in interactive questions.					
10	When I cannot solve a question in interactive questions, I stop watching the video lesson.					
11	Learning the subject from video lessons is boring.					
12	When I do the activities at the end of the video lesson, I also realize which subjects I lack knowledge about.					
13	I would like to see my name at the top of the points ranking of the activities at the end of the video lesson.					
14	The activities at the end of the video lesson help me to reinforce the topic.					
15	The activities at the end of the video lesson increase the retention of my learning.					
16	I do not continue the activity at the end of the video lesson when I cannot meet the time limit.					
17	I always search for the correct solutions to the questions I cannot solve in the activities at the end of the video lesson.					
18	I watched the video lesson several times before coming to the face-to-face lesson.					
19	I only watch video lessons because my teacher gives me homework.					

(continued)

	Evaluation of Teaching Conducted with Video Lessons (ETCVL) subscale	Strongly Disagree	Disagree	Neither agree nor disagree	Agree	Strongly Agree
20	I can watch the video lessons over and over again without getting bored.					
21	It allows me to determine the most appropriate time for my learning environment.					
22	It increases my level of active participation in the lesson.					
23	It allows me to manage my own learning process.					
24	I do not find it safe to use unfamiliar learning platforms.					
25	I do not want to provide my personal information to learning platforms, even for educational purposes.					

	Evaluation of Teaching Conducted with Face-to-Face Lessons (ETCFFL) Subscale	Strongly Disagree	Disagree	Neither agree nor disagree	Agree	Strongly Agree
1	At the beginning of a face-to-face lesson, my teacher summarizes the topic, which helps me to remember what I have learned.					
2	It is a waste of time for my teacher to summarize what I have learned in video lessons in face-to-face lessons.					
3	I get bored when my teacher summarizes the topic in video lessons at the beginning of a face-to-face lesson.					
4	I learn better when I attend a face-to-face lesson by watching a video lesson.					
5	I feel more confident when I attend a face-to-face lesson by watching a video lesson.					
6	I would like to participate more in face-to-face lessons as I have knowledge about the subject by watching videos.					
7	I would like to be the first person to answer questions when I attend a face-to-face class by watching video lessons.					