2024

Volume 17, Issue 1
Cilt 17, Sayı 1

Kuramsal
Eğitimbilim Dergisi
JOURNAL OF THEORETICAL EDUCATIONAL SCIENCE

Afyon Kocatepe Üniversitesi
Eğitim Fakültesi

ISSN: 1308-1659

http://www.keg.aku.edu.tr

Kuramsal Eğitimbilim

KURAMSAL EĞİTİMBİLİM DERGİSİ*
Journal of Theoretical Educational Science
ISSN: 1308-1659

Publisher
Afyon Kocatepe University

Owner
Dr. Nusret KOCA (Dean) on behalf of Afyon Kocatepe University

* **Journal of Theoretical Educational Science** is a quarterly peer-reviewed online journal.

Abstracting and Indexing
ERIC, ULAKBİM TR Dizin, EBSCO, ERIH Plus, DOAJ, Türk Eğitim İndeksi (TEİ), SOBIAD

Redactions
Res. Asst. Merve VEZİR

English Language Editor
Res. Asst. Gamze KAPUCU

Address
Afyon Kocatepe Üniversitesi, Eğitim Fakültesi, A.N.S. Kampüsü, 03200 Afyonkarahisar, Turkey
Tel: +90 272 2281418
e-mail: editorkebd@gmail.com

CONTENTS

FROM THE EDITOR

Dear Colleagues,

The *Journal of Theoretical Educational Science* is happy to publish the first issue of 2024! In this issue, you will find ten research articles by 20 authors. We are glad that these articles represent the different disciplines of education.

We should also express our sincere thanks to the Editorial Board, reviewers, and authors for their invaluable contributions. We look forward to receiving submissions from different parts of the world for the following issues!

Kindest regards,

Fatih GÜNGÖR, PhD
Afyon Kocatepe University
Faculty of Education

Journal of Theoretical Educational Science, 17(1), 1-27, January 2024
Kuramsal Eğitimbilim Dergisi, 17(1), 1-27, Ocak 2024
[Online]: http://dergipark.org.tr/akukeg
DOI number: http://doi.org/10.30831/akukeg.1207350

Cross-National Measurement of Mathematics Intrinsic Motivation: An Investigation of Measurement Invariance with MG-CFA and the Alignment Method Across Fourteen Countries

Matematikte İçsel Motivasyonun Ülkeler Arası Ölçümü: On Dört Ülkede MG-CFA ve Hizalama Yöntemi ile Ölçme Değişmezliğinin İncelenmesi

Mahmut Sami YİĞİTER*

Received: 19 November 2022 **Research Article** **Accepted:** 27 November 2023

ABSTRACT: One of the main objectives of international large-scale assessments is to make comparisons between different countries, education policies, education systems, or subgroups. One of the main criteria for making comparisons between different groups is to ensure measurement invariance. The purpose of this study was to test the measurement invariance of the mathematics intrinsic motivation scale across 14 countries. For this purpose, the "students like learning mathematics" scale, which measures intrinsic motivation for mathematics, was included in the TIMSS 2019 cycle. The study sample consisted of a total of 152992 students, 70192 4th grade and 82800 8th grade students from 14 different countries participating in the TIMSS 2019 cycle. Measurement invariance was tested with Multi-Group Confirmatory Factor Analysis (MG-CFA) and Alignment Method. The mathematics intrinsic motivation scale provides only configural invariance according to MG-CFA at the 4th grade level, whereas the scale provides approximate invariance according to the alignment method. At the 8th grade level, the scale provides configural and metric invariance according to MG-CFA, whereas the scale provides approximate invariance according to the alignment method. The results indicate that the mathematics intrinsic motivation scale provides approximate measurement invariance at both grade levels and that comparisons can be made between the scores of the identified countries.

Keywords: Mathematics, motivation, intrinsic motivation, measurement invariance, cross-national difference, students like learning mathematics, alignment method.

ÖZ: Geniş ölçekli uluslararası değerlendirmelerin temel amaçlarından biri, farklı ülkeler, eğitim politikaları, eğitim sistemleri veya alt gruplar arasında karşılaştırmalar yapmaktır. Farklı gruplar arasında karşılaştırma yapmanın temel ölçütlerinden biri de ölçme değişmezliğinin sağlanmasıdır. Bu çalışmanın amacı, matematik içsel motivasyon ölçeğinin 14 ülke arasında ölçme değişmezliğini test etmektir. Bu amaçla, matematiğe yönelik içsel motivasyonu ölçen "öğrenciler matematik öğrenmeyi sever" ölçeği TIMSS 2019 döngüsüne dahil edilmiştir. Çalışmanın örneklemi TIMSS 2019 döngüsüne katılan 14 farklı ülkeden 70192 4. sınıf ve 82800 8. sınıf öğrencisi olmak üzere toplam 152992 öğrenciden oluşmaktadır. Ölçme değişmezliği, Çok Gruplu Doğrulayıcı Faktör Analizi (MG-CFA) ve Hizalama Yöntemi ile test edilmiştir. Matematik içsel motivasyon ölçeği, 4. sınıf düzeyinde MG-CFA'ya göre sadece yapısal değişmezliği sağlarken, hizalama yöntemine göre yaklaşık değişmezliği sağlamaktadır. 8. sınıf düzeyinde ise ölçek, MG-CFA'ya göre konfigüral ve metrik değişmezliği sağlarken, hizalama yöntemine göre yaklaşık değişmezliği sağlamaktadır. Sonuçlar, matematik içsel motivasyon ölçeğinin her iki sınıf düzeyinde de yaklaşık ölçme değişmezliğini sağladığını ve belirlenen ülkelerin puanları arasında karşılaştırmalar yapılabileceğini göstermektedir.

Anahtar kelimeler: Matematik, motivasyon, içsel motivasyon, ölçme değişmezliği, ülkeler arası farklılık, öğrenciler matematik öğrenmeyi sever, hizalama yöntemi.

* Inst. MSc., Social Sciences University of Ankara, Ankara, Turkey, mahmutsamiyigiter@gmail.com, https://orcid.org/0000-0002-2896-0201

Citation Information
Yiğiter, M. S. (2024). Cross-national measurement of mathematics intrinsic motivation: An investigation of measurement invariance with MG-CFA and alignment method across fourteen countries. *Kuramsal Eğitimbilim Dergisi [Journal of Theoretical Educational Science]*, *17*(1), 1-27.

ISSN: 1308-1659

The basis of effective mathematics teaching is to support positive attitudes towards learning mathematics and to encourage learning mathematics. According to social cognitive theory, an individual's self-efficacy in a subject affects his/her motivation in that subject (Schunk & DiBenedetto, 2020). Then, academic success emerges with the effect of motivation on performance (Yıldırım, 2011). An individual's motivation to accomplish a task enables him/her to orient himself/herself towards that field and to work persistently in that field. Many studies have shown that student motivation and academic achievement are related (Ahmed et al., 2010; Cleary & Chen, 2009; İlter, 2021; Woolley et al., 2010).

International Large-Scale Assessment (ILSA) programmes have been started to be carried out with the participation of many countries since the end of the 20th century in order to make comparisons between the education systems of countries and to determine student achievement (Cardoso, 2020). Trends in International Mathematics and Science Study (TIMSS) is an assessment study that evaluates the academic achievement of 4th and 8th-grade students in the fields of mathematics and science skills every four years and monitors the achievement differences between countries over time and the results of countries' attempts to increase the level of achievement (Mullis & Martin, 2017). In addition to measuring mathematics and science skills, TIMSS includes many scales and scales that measure cognitive and affective characteristics of students and teachers, such as self-confidence, motivation, school belonging, and peer bullying (Yin & Fishbein, 2019). Students' affective characteristics for mathematics achievement have been measured since 1995. One of the scales included in the TIMSS 2019 student survey is the "Students Like Learning Mathematics" scale, which measures mathematics intrinsic motivation. There are many studies reporting that intrinsic motivation has positive and significant effects on mathematics achievement (Akben-Selcuk, 2017; Guo et al., 2015; Hooper et al., 2020). Therefore, it is important that the scales measuring the characteristics of the mathematics intrinsic motivation scale perform valid and reliable measurements in order to accurately and objectively reveal the relationships between intrinsic motivation and mathematics achievement in cross-country and cross-cultural comparisons.

One of the main goals of ILSAs is to make comparisons between different countries, education systems, subgroups, and individuals (Engel & Rutkowski, 2021). There is increasing methodological discussion about the use of data from ILSAs for making comparisons (Gustafsson, 2018; Rutkowski & Svetina, 2014). One of the basic and critical criteria for making comparisons between different groups is the establishment of measurement invariance (Putnick & Bornstein, 2016). Measurement invariance is a statistical property that analyses whether the scale has equivalent psychometric values between the different groups or sub-groups to which it is applied (Raykov, 2004). A measurement tool should measure the construct in a psychometrically equivalent way in each subgroup. If the psychometric properties of the measurements obtained from subgroups differ, it would not be correct to generalize the results (Başusta & Gelbal, 2015). Therefore, a measurement tool should measure the construct equivalently in each subgroup. With measurement invariance, showing that the factor loadings, inter-dimensional correlations, and error variances of a scale are the same in each group will show that the measurement tool has an equivalent structure in different groups (Jöreskog & Sörbom, 1993). Researchers obtain evidence on whether

the scale measures the same construct in subgroups (Millsap & Olivera-Ogilar, 2012; Uyar & Doğan, 2014). Failure to provide measurement invariance is a validity problem for the measurement tool. Therefore, interpretations regarding the results of group comparisons based on such a measurement tool may also be incorrect (Vandenberg & Lance, 2000). Demonstrating that measurement invariance is established will also provide validity evidence for the measurement tool. He et al. (2019), in their study on cross-cultural comparability with TIMSS and PISA data, state that comparisons made without examining measurement invariance may lead to inaccurate results, hence the importance of testing measurement invariance.

Mathematics Intrinsic Motivation

Intrinsic motivation is defined as the enjoyment received while engaging in an activity and the drive to perform this activity (OECD, 2013). Students' willingness to learn mathematics stems from the fact that they find mathematics interesting and fun (Ryan & Deci, 2009). Intrinsically motivated students take action to "learn" rather than doing something for a purpose. Extrinsically motivated students see the task as a means to achieve a goal and complete this task with concepts such as "grade, reward, competition, performance, external evaluation." For example, a student who studies for an exam to get a good grade is motivated (extrinsically) by the grade he/she gets. Students who study only because they enjoy it are intrinsically motivated. Previous research has revealed that students' intrinsic motivation can be positively influenced by the teacher's approach, curriculum, methods applied in the lesson, learning environment design, and practices (Freiberger et al., 2012; Middleton, 1995; Weidinger et al., 2017). Mueller et al. (2011) state that solving open-ended questions and discussing in the classroom environment can increase intrinsic motivation. Similarly, problem-based learning increases students' intrinsic motivation (Henderson & Landesman, 1995). There are studies indicating that applied learning and active learning activities increase intrinsic motivation (Barak & Asad, 2012; Nugent et al., 2010). In addition, when students accomplish a difficult task, their motivation increases (Middleton, 1995).

Intrinsic motivation is a source of energy and a precondition for behaviour (Malone & Lepper, 2021). Previous studies show that there is a strong relationship between intrinsic motivation and mathematics achievement (Hooper et al., 2020; İlhan & Çetin, 2013; Mullis et al., 2017; Tavani & Losh, 2003; Zembat et al., 2018). Intrinsic Motivation for Mathematics is measured with nine items under the "Students Like Learning Mathematics" scale at both 4th and 8th-grade levels in TIMSS 2019.

Measurement Invariance with Multi-Group Confirmatory Factor Analysis

Measurement invariance investigates whether a latent construct is measured consistently across different groups, categories, or times (Cheung & Rensvold, 2002; Sözer et al., 2021). In other words, if the psychometric values obtained from the same scale differ in different groups, it can be said that measurement invariance cannot be provided. Individuals in different groups who are equivalent to each other in terms of the measured feature (construct) are expected to get the same observed score from a test. If the individuals are the same in terms of the construct measured, but their observed scores are different, it can be said that measurement invariance of the scale (test) cannot be ensured (Schmitt & Kuljanin, 2008). If measurement invariance cannot be proved, it is not correct to interpret the results of intergroup comparisons. The reason

for this is that it cannot be known whether the difference between the groups is due to a real construct difference or the difference between the responses to the scale items (Adibatmaz & Yildiz, 2020; Horn & McArdle, 1992). Therefore, it is important to test for measurement invariance before making inferences about measurements from two or more groups (Yiğiter, 2023).

One of the most commonly used methods in the literature to test measurement invariance is Confirmatory Factor Analysis (CFA) (Schmitt & Kuljanin, 2008; Van De Schoot et al., 2015). Measurement invariance is tested in four hierarchical stages with the Multi-Group CFA (MG-CFA) method. These stages are configural invariance, metric invariance, scalar invariance, and strict invariance (Meredith, 1993).

Configural Invariance

It is the first hierarchical stage of measurement invariance. At this stage, whether the groups have the same factor structure is tested. For this purpose, the equivalence of factors and the pattern of factor loadings are analysed at this stage (Taris et al., 1998). No parameter restriction is made at this stage. If configural invariance is provided, it can be stated that the groups measure the same construct (Wu et al., 2007). If configural invariance is not provided, it is stated that the groups measure different constructs, and the further stages of measurement invariance are not passed. Configural invariance is also referred to as structural invariance in the literature.

Metric Invariance

When it is shown that configural invariance is provided, a metric invariance test can be performed (Milfont & Fischer, 2010). In metric invariance, the equality of factor loadings in different groups is tested. In other words, factor loadings estimated from one group are fixed to the other group, and the fit indices of the model are examined. If metric invariance is provided, comparisons between groups based on factor loadings can be defended (Gregorich, 2006). Metric invariance is also known as weak invariance (Meredith, 1993).

Scalar Invariance

If metric invariance is provided, the scalar invariance stage is proceeded. In this stage, the equivalence of both factor loadings and regression constants between groups is tested. In other words, at this stage, where the equivalence of factor variance and covariances between groups is tested, the equivalence of factor loadings is also examined. If scalar invariance is provided, it means that the means and factor loadings of the observed variables can be compared (Gregorich, 2006). Scalar invariance is also known as strong invariance.

Strict Invariance

At this stage, which is the last step of measurement invariance, in addition to the restrictions in the previous stages, the equality of error variances is also tested (Vandenberg & Lance, 2000). Scales that claim to measure the same construct across groups should provide strict invariance. By ensuring strict invariance, measurement invariance will be fully provided.

Measurement Invariance with Multi-Group Alignment Method

In measurement invariance with MG-CFA - especially when the number of groups is large - it becomes difficult to ensure model fit across stages. MG-CFA assumes strict invariance, which may be an unreachable goal when the number of groups increases. When measurement invariance is rejected in MG-CFA, partial models with free estimation of some item parameters can be tested, but there is no guarantee that these models will also provide measurement invariance (Asparouhov & Muthén, 2014). On the other hand, when measurement invariance is rejected with MG-CFA, the causes of invariance are not properly identified by the analysis. Moreover, the probability of incorrect calculation will increase as a result of MG-CFA's pair-by-pair comparison of groups. Therefore, MG-CFA is not practical when comparing the measurement invariance of a large number of groups (Sırgancı et al., 2020).

A more recent approach, the Alignment method, greatly simplifies the measurement invariance analysis. It allows testing the invariance of parameters according to items and groups. In other words, in the alignment method, it can be determined which group contributes to measurement invariance. The alignment method proposed by Asparouhov and Muthén (2014) provides a result that minimises parameter invariance between groups in an iterative process, similar to rotation in exploratory factor analysis (Glassow et al., 2021).

Aim and Significance of the Study

The purpose of this study is to examine the measurement invariance of the mathematics intrinsic motivation scale in the TIMSS 2019 cycle according to 14 countries. When the invariance studies in the literature were examined, it was seen that measurement invariance studies were carried out according to culture, region, language, and gender (Alatlı, 2020; Bağdu Söyler et al., 2021; Ertürk & Erdinç-Akan, 2018; Uyar & Doğan, 2014;). Studies examining measurement invariance according to the mathematics intrinsic motivation scale are quite limited (He et al., 2019). No study was found that tested the measurement invariance of the mathematics intrinsic motivation scale across cultures with the TIMSS 2019 data. This study is important both because it focuses on the cross-cultural measurement invariance of the mathematics intrinsic motivation scale and because it has not been investigated before.

Related Studies

Glasgow et al. (2019) examined the measurement invariance of mathematics teachers' Job satisfaction, School emphasis on academic success, School condition and resources, Safe and orderly school, and teacher Self-efficacy scales obtained from TIMMS 2015 teacher surveys across 46 countries. The results show that only three constructs provide metric invariance. The results of measurement invariance with the Alignment Optimisation method show that all five constructs provide approximate invariance so that these constructs can be validly compared across educational systems.

Sırgancı et al. (2020), in their study explaining the basic concepts and processes of the alignment method, compared the measurement invariance of 56 countries on the Instrumental Motivation Scale data in the PISA 2015 cycle with both MG-CFA and alignment method. MG-CFA findings show that the scale provides only configural invariance. Then, the measurement invariance findings with the alignment method

provide more detailed information about which countries and which items contribute better to measurement invariance.

Tekin and Cobanoglu-Aktan (2021) examined the measurement invariance of collaborative problem-solving skills in the PISA 2015 cycle between Singapore, Norway, and Turkey with the MG-CFA method. The results of the study show that the construct provides only configural invariance but not metric, scalar, and strict invariance.

Method

This study is descriptive research since it aims to determine whether the "Students Like Learning Mathematics" scale from TIMSS 2019, which is included in the TIMSS 2019 student questionnaire and measures mathematics intrinsic motivation, does not change according to 14 different countries (Büyüköztürk et al., 2017).

Population and Sample

More than 580,000 students from 64 countries around the world participated in TIMSS 2019, which was administered by the International Association for the Evaluation of Educational Achievement (IEA). Approximately 330.000 of these students are in Grade 4, and 250.000 are in Grade 8. The sample of this study consists of a total of 152.992 students, 70.192 of whom are 4th graders and 82.800 of whom are 8th graders, who participated in TIMSS 2019 from 14 countries. In determining the sample, countries were determined by taking into account the differentiation in terms of language, continent, culture, and achievement rankings. In addition, since the research was conducted for both the 4th and 8th-grade levels, the countries that participated in TIMSS 2019 at both grade levels were selected for the current study. The distribution of the sampled countries according to their sample sizes, languages, and continents is given in Table 1.

Table 1

Countries in the Sample and Their Characteristics

Country Name	ISO Country Code	Language	Continent	4th Grade	8th Grade
Australia	36	English	Oceania	5664	8898
Chile	152	Spanish	South America	4039	4061
Hungary	348	Hungarian	Europe	4433	4537
Italy	380	Italian	Europe	3666	3600
Japan	392	Japanese	Asia	4162	4443
Morocco	504	Moroccan Arabic and others	Africa	7645	8440
Portugal	620	Portuguese	Europe	4256	3348
Russia	643	Russian	Europe-Asia	3993	3890
Saudi Arabia	682	Arabic	Asia	5334	5634
Singapore	702	English, Malay, and others	Asia	4362	4165
South Africa	710	Afrikaans	Africa	11729	20717

Sweden	752	Swedish	Europe	3816	3907
Turkey	792	Turkish	Asia	3998	3978
Pakistan	926	Urdu	Asia	3095	3182
Total				70192	82800

As seen in Table 1, 14 countries in the sample are located on five different continents. In addition, the official languages of each of these countries are different.

Data Source

The data were obtained from the database at "https://timss2019.org/international-database/". The TIMSS administration includes mathematics and science achievement tests as well as student, teacher, school, and home questionnaires. In addition, there are also items that examine the affective characteristics of students in mathematics and science. This study was limited to the "Students Like Learning Mathematics" scale, which measures intrinsic motivation for mathematics. The items and codes in this scale are given in Table 2.

Table 2

Items and Codes in the Scale

Item Code		Description
4. Grade	8. Grade	
ASBM02A	BSBM16A	I enjoy learning mathematics
ASBM02B	BSBM16B	I wish I did not have to study mathematics [R]
ASBM02C	BSBM16C	Mathematics is boring [R]
ASBM02D	BSBM16D	I learn many interesting things in mathematics
ASBM02E	BSBM16E	I like mathematics
ASBM02F	BSBM16F	I like any schoolwork that involves numbers
ASBM02G	BSBM16G	I like to solve mathematics problems
ASBM02H	BSBM16H	I look forward to mathematics lessons
ASBM02I	BSBM16I	Mathematics is one of my favourite subjects

There are nine items in this scale, as seen in Table 2. Items 2 and 3 are reverse-coded questions. All items were scored on a 4-point Likert rating scale with the options "agree a lot, agree a little, disagree a little, disagree a lot."

Data Analysis

All analyses in this study were performed with the open-source R program. "dplyr" (Wickham et al., 2019) for data manipulation, "lavaan" (Rosseel, 2012) for CFA and MG-CFA analyses, "sirt" (Robitzsch, 2019) for alignment method, "naniar" (Tierney et al., 2021) for missing data analysis, "mvdalab" (Afanador et al., 2016) for missing data imputation, "Performance Analytics" (Peterson et al., 2018) for normality analysis. Grade 4 and Grade 8 data obtained from the TIMSS 2019 database were

analysed separately. Before starting the data analysis, missing data, outliers, normality, and multicollinearity were examined.

4th Grade Level

It is seen that there are a total of 73336 students participating in the exam at the 4th grade level from 14 countries. Firstly, missing data, which are questions not answered by the students, were analysed. It was observed that the rate of missing data on a variable basis varied between 2.5% (1878/73336) and 4.8% (3546/73336). On the basis of all data, the rate of missing data reaches 11.3% (8323/73336). Little MCAR test was performed to examine the randomness of the missing data. The Little MCAR test results show that the missing data is random and does not contain any pattern (LittleMCAR=5091; df=2604; p>0.05). Since the missing data is more than 10%, missing data assignment was preferred instead of the listwise deletion method in order not to lose the analysis power. Before assigning missing data, 1065 participants who left the entire scale used in the study blank were deleted. Then, the missing data assignment was performed with the EM algorithm. The extreme value analysis was performed in two stages. Before the extreme value analysis, participants who answered carelessly were identified. Participants who give the same responses to questions with reverse coding and questions with normal coding in the scale exhibit careless responding behaviour (Woods, 2006). It is stated that these participants have disruptive effects on the factor structure (Kam, 2019). In the 4th grade data, 1560 participants with careless responding behaviour were identified and excluded from the sample (Kam & Meyer, 2015). In the second stage, total and standardised Z scores were calculated according to the scales. According to the z score, participants who were outside the [-3,+3] range were determined as outliers (Kaliyaperumal et al., 2015). According to the Z scores, 519 observations outside this range were identified as outliers and deleted from the data. Skewness and kurtosis values of the variables were analysed to determine whether the data were normally distributed. Since the kurtosis and skewness coefficients were in the range of [-1.5,+1.5], it was decided that the data were normally distributed (Tabachnick & Fidell, 2013). In order to examine the multicollinearity, the VIF (variance inflation factor) value was calculated for all items. The highest VIF value was found to be 3.55 in the item coded BSBM16E. It was decided that there was no multicollinearity problem since a multicollinearity problem would occur if the VIF value was greater than 5 (Kline, 2011).

8th Grade Level

It is seen that there are 84345 students from 14 countries who participated in the exam at the 8th-grade level. Firstly, missing data, which are questions not answered by the students, were analysed. It was observed that the missing data rates on a variable basis varied between 1.5% (1275/84345) and 3.7% (3142/84345). On the basis of all data, the missing data rate reaches 8.21% (6927/84345). Little MCAR test was performed to examine the randomness of missing data. The Little MCAR test results show that the missing data is random and does not contain any pattern (LittleMCAR=3.179; sd=1701; p>0.05). Before assigning missing data, 860 participants who left the entire scale used in the study blank were deleted. Missing data assignment was made with the EM algorithm. The extreme value analysis was performed in two

stages. Before the extreme value analysis, participants who carelessly responded were identified. In this data, 685 participants who responded carelessly were identified and removed from the sample (Kam & Meyer, 2015). Since there was no value outside the range of [-3,+3] according to Z scores, it was decided that there was no outlier (Kaliyaperumal et al. , 2015). The fact that the kurtosis and skewness coefficients are in the range of [-1.5,+1.5] indicates that the data are normally distributed (Tabachnick & Fidell, 2013). The highest VIF value calculated to examine the multicollinearity problem was found to be 4.37 in the item coded BSBM16E. It was decided that there was no multicollinearity problem since a multicollinearity problem would occur if the VIF value was greater than 5 (Kline, 2011).

Multi-Group CFA analyses were performed using the lavaan package in R (Rosseel, 2012). Another reason why this package was preferred is that it allows the use of sample weights when estimating with the MG-CFA model. In large-scale assessments, all participants in the population (all students at the relevant grade level) cannot be included in the sample due to time and financial limitations. In order to overcome this limitation and to ensure the generalisability of the sample to the population, sampling weights are used (Arıkan et al., 2020). Student weights in the TIMSS 2019 data were added to the model as sampling weights. In this study, country code as a categorical variable and items of the intrinsic motivation scale as ordinal variables were used. It is recommended to use WLS (weighted least squares), WLSMV (robust weighted least squares), or ULS (unweighted least squares) methods that are robust to violations of assumptions as estimation methods (Brown, 2006; Koğar & Yılmaz Koğar, 2015). Therefore, the WLSMV estimation method, which is reported to give good results in the MG-CFA model, was used as the estimation method (Forero et al., 2009).

Measurement invariance was analysed by testing four hierarchical stages with MG-CFA. These four stages are configural invariance, metric invariance, scalar invariance, and strict invariance (Vandenberg & Lance, 2000). To examine the model-data fit between the stages, χ^2, χ^2/df, $RMSEA$, $SRMR$, TLI, CFI and ΔCFI values were reported. Acceptable levels of these values are presented in Table 3 (Hu & Bentler, 1999).

Table 3

Acceptable Ranges of Goodness of Fit Indices

Fit Indices	Acceptable Fit	Good Fit
χ^2/df	$3 < \chi^2/df < 5$	$0 < \chi^2/df < 3$
CFI	0.95 < CFI < 0.97	0.97 < CFI < 1
TLI	0.95 < TLI < 0.97	0.97 < TLI < 1
RMSEA	0.05 < RMSEA < 0.08	0.00 < RMSEA < 0.05
SRMR	0.05 < SRMR < 0.08	0.00 < SRMR < 0.05

Note. (Hu & Bentler, 1999)

In the MG-CFA method, once it was determined that the fit at a particular stage was satisfactory, the analysis proceeded to the next stage. There are studies suggesting

that the significance can be tested according to the difference of chi-square values ($\Delta\chi^2$) in determining whether inter-stage invariance is achieved (Schmitt & Kuljanin, 2008). However, the chi-square difference test rejects the null hypothesis with too much power as the sample size increases. Therefore, Cheung and Rensvold (2002) suggested examining the change in CFI value (ΔCFI) as an alternative to $\Delta\chi^2$. In this study, a difference of ΔCFI less than or equal to 0.01 was used as a criterion to determine whether inter-stage invariance was achieved (Cheung & Rensvold, 2002).

In measurement invariance with the Alignment Method, the fit of a configural model without restriction between groups was first assessed. Then, it was optimised with a component loss function to minimise the invariance between the means of each factor and the variances of the groups under the configural model (Asparouhov & Muthén, 2014). The tolerance criteria proposed by Robitzsch (2020) were used (factor loadings (λ=.40) and intercepts (ν=.20)). The alignment strength for the parameters was determined as .25 (Fischer & Karl, 2019). The equivalence of the parameters was interpreted with the R^2 value. R^2 values close to 1 indicate that there is more invariance (Asparouhov & Muthén, 2014). In determining the measurement invariance, the cut-off criterion of 25% of the invariant parameter ratio of the intercept and slope parameters was used (Asparouhov & Muthén, 2014).

Results

In this section, the findings obtained from the mathematics intrinsic motivation scale for 4th and 8th grade levels according to TIMSS 2019 data are presented. Analyses were conducted separately for both grade levels. Firstly, CFA was conducted. Finally, measurement invariance was tested with MG-CFA.

CFA Results

CFA analysis was performed to check the unidimensional factor structure. The results of the CFA analyses are presented under separate subheadings according to the grade level.

CFA Results (4th Grade Level)

The fit indices obtained from the CFA analysis are presented in Table 4.

Table 4

Fit Statistics of the CFA Model (4th Grade)

χ^2	df	χ^2/df	RMSEA	*SRMR*	*TLI*	*CFI*
14905.4	27	552.0	0.049	0.048	0.985	0.989

The χ^2/df value is expected to be less than 5 to ensure model-data fit. However, since the χ^2/df ratio is especially affected by the sample size, it is recommended to use other indexes in model-data fit. The results indicate an RMSEA of 0.049, an SRMR of 0.048, a CFI index of 0.989, and a TLI index of 0.985. According to these values, it can be concluded that the model-data fit is at an acceptable level (Schreiber et al., 2006).

The standardised factor loadings, AVE (average variance extracted) values, and reliability coefficients obtained from the CFA model are presented in Table 5.

Table 5

Standardised Factor Loadings, AVE Values, and Reliability Coefficients

Items	Factor Loadings	AVE	Cronbach Alfa	McDonald's Omega
ASBM02A	0.689			
ASBM02B	0.483			
ASBM02C	0.577			
ASBM02D	0.504			
ASBM02E	0.802	%52.55	0.902	0.905
ASBM02F	0.679			
ASBM02G	0.726			
ASBM02H	0.827			
ASBM02I	0.876			

According to the CFA results, factor loadings are expected to be 0.30 and above (Harrington, 2009). These results show that all of the factor loadings are at a sufficient level. The AVE (average variance extracted) values in the Table 5 are calculated from the factor loadings and give information about what percentage of the variance of the items the factor explains. Generally, this value is desired to be 50% and above (Hair et al., 2014). It can be said that the scale has a good AVE value. On the other hand, internal consistency coefficients provide evidence of whether a homogeneous structure is measured in addition to providing evidence of reliability (Aybek, 2022). Cronbach's Alpha and McDonald's Omega coefficients of the scales are presented in Table 5 on internal consistency. According to Büyüköztürk (2011), a reliability coefficient of 0.70 and above is considered sufficient for scale reliability. Kline (2011) categorises 0.90 and above as excellent, 0.80 and above as good, and 0.70 and above as acceptable. As seen in the Table 5, both Cronbach's Alpha and McDonald's Omega coefficients are more than 0.70 and are considered to be at a good level.

The results of the separately conducted CFAs for each country are presented in Appendix 1. It is observed that the fit indices for Morocco, South Africa, and Saudi Arabia are notably low. In addition, it is noticeable that the factor loadings of the reverse-coded items in these countries are quite low. It can be stated that the CFA results of other countries are at a good level in terms of fit indices, factor loadings, AVE, and reliability values.

CFA Results (8th Grade Level)

The fit indices obtained from the CFA analysis are presented in Table 6.

Table 6

Fit Statistics of the CFA Model (8th Grade)

χ^2	df	χ^2/df	RMSEA	$SRMR$	TLI	CFI
22238.6	27	823.6	0.048	0.039	0.993	0.994

Table 6 shows that RMSEA is 0.048, SRMR is 0.039, CFI index is 0.994, and TLI index is 0.993. According to these values, it can be said that the model-data fit is at a good level (Schreiber et al., 2006).

The standardised factor loadings, AVE (average variance extracted) values, and reliability coefficients obtained from the CFA model are presented in Table 7.

Table 7

Standardised Factor Loadings, AVE Values, and Reliability Coefficients

Items	Factor Loadings	AVE	Cronbach Alfa	McDonald's Omega
BSBM16A	0.796			
BSBM16B	0.597			
BSBM16C	0.677			
BSBM16D	0.644			
BSBM16E	0.929	%60.60	0.928	0.931
BSBM16F	0.780			
BSBM16G	0.840			
BSBM16H	0.820			
BSBM16I	0.958			

When Table 7 is analysed, it is seen that all of the factor loadings are at a sufficient level. It can be said that the AVE value is at a good level. In addition, both Cronbach's Alpha and McDonald's Omega coefficients are more than 0.70 and are accepted to be at a good level.

CFA results separately by country are presented in Appendix 2. It can be stated that the CFA fit indices of all countries are at a good or acceptable level.

Measurement Invariance Results

Measurement invariance was analysed in two stages. Firstly, under the known MG-CFA model, configurational invariance, metric invariance, scalar invariance, and strict invariance were tested in hierarchical order. In the second stage, measurement invariance was analysed by the alignment method. In the MG-CFA model, the ΔCFI

value was taken into account in determining whether the invariance was achieved between two hierarchical stages. When this value is ΔCFI<0.01 between the stages, it is interpreted that measurement invariance is provided at the relevant stage (Cheung & Rensvold, 2002). In this section, the measurement invariance of the mathematics intrinsic motivation scale according to the 4th and 8th-grade data was carried out and reported separately according to the grade level.

Measurement Invariance Results (4th Grade)

The measurement invariance findings of the mathematics intrinsic motivation scale with the MG-CFA method according to 14 countries are presented in Table 8.

Table 8

Measurement Invariance by Countries (4th Grade)

Stage	χ^2	df	χ^2/df	$RMSEA$	$SRMR$	TLI	CFI	ΔCFI
Configural	5506.1	378	14.5	0.053	0.043	0.984	0.987	-
Metric	12872.1	482	26.7	0.072	0.068	0.971	0.972	0.015

When the configural invariance findings are analysed, it is evident that the scale's structure is compatible with configural invariance in all country groups ($RMSEA$ < .08, $SRMR$ < .08, TLI > .95, CFI > .95). After configural invariance, metric invariance was analysed. However, as the difference in CFI value between the configural invariance and metric invariance stages is 0.015 (ΔCFI>0.01), it is apparent that metric invariance cannot be established across countries.

According to the results of the analyses conducted with the MG-CFA method, when the ΔCFI (ΔCFI < 0.01) values and goodness of fit statistics ($RMSEA$ < .08, $SRMR$ < .08, TLI > .95, CFI > .95) of the 4th-grade data were taken into consideration, it was concluded that 14 different countries provided only configural invariance. According to this result, it can be stated that the item-factor structure is equivalent between groups, but factor loadings, variances, covariances, and error variances are not equivalent between groups.

4th-grade level Mathematics Intrinsic Motivation scale measurement invariance with alignment method findings are presented in Table 9 and Table 10.

Table 9

Each item's Alignment Results of 14 Countries

Stage	Items	Mean	SD	Min	Max	R^2	%
Factor Loadings	ASBM02A	0.64	0.05	0.52	0.72	0.991	%0.0
	ASBM02B	0.50	0.12	0.30	0.67		
	ASBM02C	0.59	0.11	0.38	0.73		
	ASBM02D	0.47	0.09	0.36	0.63		
	ASBM02E	0.78	0.03	0.74	0.82		

	ASBM02F	0.61	0.05	0.50	0.72		
	ASBM02G	0.69	0.05	0.60	0.75		
	ASBM02H	0.79	0.03	0.74	0.84		
	ASBM02I	0.85	0.03	0.79	0.88		
Intercept	ASBM02A	1.66	0.07	1.56	1.79	0.995	%8.7
	ASBM02B	1.85	0.20	1.46	2.30		
	ASBM02C	1.84	0.14	1.54	2.08		
	ASBM02D	1.58	0.10	1.43	1.73		
	ASBM02E	1.72	0.03	1.67	1.80		
	ASBM02F	1.85	0.15	1.64	2.15		
	ASBM02G	1.89	0.09	1.80	2.12		
	ASBM02H	2.00	0.08	1.82	2.14		
	ASBM02I	2.02	0.07	1.89	2.17		

The alignment method was invariant for the factor loadings (R^2=.991) and the intercepts of the items (R^2=.995) of the Mathematics Intrinsic Motivation Scale, as seen in Table 9. While there is no non-invariant parameter in the factor parameters (0.0%), 11 parameters are non-invariant according to the item intercepts (8.7%). Since the percentage of non-invariant parameters is lower than 25%, it can be stated that the scale provides approximate measurement invariance (Asparouhov & Muthén, 2014).

Table 10 presents the invariance findings of intercept and factor loading parameters on an item and country basis. Countries in brackets and bold are labelled as non-invariant since they show more variability than the tolerance parameter.

Table 10

Each Item's Alignment Results of 14 Countries

Items	Intercepts	Factor Loadings
ASBM02A	36, 152, 348, 380, 392, 504, 620, 643, 682, 702, 710, 752, 792, 926	36, 152, 348, 380, 392, 504, 620, 643, 682, 702, 710, 752, 792, 926
ASBM02B	36, (**152**), 348, 380, (**392**), (**504**), **(620), (643)**, 682, 702, 710, 752, 792, 926	36, 152, 348, 380, 392, 504, 620, 643, 682, 702, 710, 752, 792, 926
ASBM02C	36, 152, 348, 380, 392, (**504**), **(620),** 643, 682, 702, 710, 752, 792, 926	36, 152, 348, 380, 392, 504, 620, 643, 682, 702, 710, 752, 792, 926
ASBM02D	36, 152, 348, 380, 392, 504, 620, 643, 682, 702, 710, 752, 792, 926	36, 152, 348, 380, 392, 504, 620, 643, 682, 702, 710, 752, 792, 926
ASBM02E	36, 152, 348, 380, 392, 504, 620, 643, 682, 702, 710, 752, 792, 926	36, 152, 348, 380, 392, 504, 620, 643, 682, 702, 710, 752, 792, 926

ASBM02F	**(36)**, 152, 348, 380, 392, 504, 620, 643, **(682)**, 702, 710, 752, 792, 926	36, 152, 348, 380, 392, 504, 620, 643, 682, 702, 710, 752, 792, 926
ASBM02G	**(36)**, 152, 348, 380, 392, 504, 620, 643, 682, 702, 710, 752, 792, 926	36, 152, 348, 380, 392, 504, 620, 643, 682, 702, 710, 752, 792, 926
ASBM02H	36, 152, 348, 380, 392, 504, 620, 643, 682, 702, (**710)**, 752, 792, 926	36, 152, 348, 380, 392, 504, 620, 643, 682, 702, 710, 752, 792, 926
ASBM02I	36, 152, 348, 380, 392, 504, 620, 643, 682, 702, 710, 752, 792, 926	36, 152, 348, 380, 392, 504, 620, 643, 682, 702, 710, 752, 792, 926

Note. 36: Australia, 152: Chile, 348: Hungary, 380: Italy, 392: Japan, 504: Morocco, 620: Portugal, 643: Russia, 682: Saudi Arabia, 702: Singapore, 710: South Africa, 752: Sweden, 792: Turkey, 926: Pakistan.

When Table 10 is examined, it is seen that all factor loadings of Mathematics Intrinsic Motivation are invariant in 14 countries at the 4th-grade level. According to the intercept coefficients, it can be stated that five parameters in item ASBM02B, two parameters in item ASBM02C, two parameters in item ASBM02F, one parameter in item ASBM02G, and one parameter in item ASBM02H are non-invariant. Findings on an item basis show that 4 out of 9 items are invariant across all countries, while in 5 items, some countries are non-invariant in the intercept coefficients. On the other hand, it is seen that the negatively rooted items ASBM02B and ASBM02C are the items that violate invariance the most. Findings by country suggest that factor loading parameters are invariant across all countries. According to the intercept parameter, Australia is non-invariant in two items: Morocco in two, Portugal in two, Japan in one, Russia in one, Saudi Arabia in one, Chile in one, and South Africa in one item. It can be stated that both factor loadings and intercept parameters of Hungary, Italy, Singapore, Sweden, Turkey, and Pakistan are equivalent. As Asparouhov and Muthén (2014) suggested, since the non-invariant parameter ratio of factor loading and intercept parameters is below 25%, it can be stated that the Mathematics Intrinsic Motivation scale provides approximate invariance and all groups can be compared.

Measurement Invariance Results (8th Grade)

The measurement invariance findings of the mathematics intrinsic motivation scale with the MG-CFA method according to 14 countries are presented in Table 11.

Table 11

Measurement Invariance by Countries (8th Grade)

Stage	χ^2	df	χ^2/df	$RMSEA$	$SRMR$	TLI	CFI	ΔCFI
Configural	5480.9	378	14.5	0.049	0.036	0.993	0.994	-
Metric	12919.7	482	26.8	0.067	0.054	0.986	0.987	0.007
Scalar	25908.5	586	44.2	0.086	0.076	0.977	0.974	0.013

When analyzing the configural invariance findings, the scale structure is found to be compatible across all country groups ($RMSEA$ < .08, $SRMR$ < .08, TLI > .95, CFI

> .95). After configural invariance, metric invariance was analysed. Since the difference in CFI value between the configural invariance and metric invariance stages (ΔCFI) was 0.008, it was decided that metric invariance was also provided according to the countries. Additionally, other fit indices are also observed to be at a satisfactory level at the metric invariance stage ($RMSEA$=0.066, $SRMR$=0.064, TLI= 0.987). After metric invariance, scalar invariance was analysed. Since the ΔCFI value between metric invariance and scalar invariance is 0.013 (ΔCFI>0.01), it is decided that scalar invariance is not provided. It should also be noted that the RMSEA value at the scalar invariance stage showed incompatibility ($RMSEA$ > .08).

According to the results of the analyses conducted with the MG-CFA method, considering the ΔCFI (ΔCFI < 0.01) values and goodness of fit statistics ($RMSEA$ < .08, $SRMR$ < .08, TLI > .95, CFI > .95) of the 8th-grade data, it was concluded that 14 different countries provided configural and metric invariance. According to this result, it can be stated that item-factor structure and factor loadings are equivalent between groups, but variances, covariances, and error variances are not equivalent between groups.

8th-grade level Mathematics Intrinsic Motivation scale measurement invariance with alignment method findings are presented in Table 12 and Table 13.

Table 12

Each Item's Alignment Results of 14 Countries

Stage	Items	Mean	SD	Min	Max	R^2	%
Factor Loadings	BSBM16A	0.78	0.05	0.69	0.85	0.995	%0.0
	BSBM16B	0.65	0.09	0.51	0.80		
	BSBM16C	0.68	0.06	0.59	0.80		
	BSBM16D	0.61	0.06	0.51	0.71		
	BSBM16E	0.91	0.01	0.88	0.91		
	BSBM16F	0.72	0.03	0.67	0.79		
	BSBM16G	0.82	0.03	0.76	0.87		
	BSBM16H	0.77	0.06	0.69	0.91		
	BSBM16I	0.93	0.06	0.80	1.04		
Intercept	BSBM16A	1.96	0.07	1.86	2.09	0.997	%8.7
	BSBM16B	2.12	0.19	1.93	2.67		
	BSBM16C	2.23	0.15	1.95	2.43		
	BSBM16D	1.97	0.08	1.86	2.17		
	BSBM16E	2.07	0.06	2.01	2.18		
	BSBM16F	2.33	0.14	2.07	2.53		
	BSBM16G	2.26	0.07	2.20	2.41		
	BSBM16H	2.48	0.14	2.15	2.66		
	BSBM16I	2.46	0.08	2.33	2.58		

The alignment method was invariant for the factor loadings (R^2=.995) and the intercepts of the items (R^2=.997) of the Mathematics Intrinsic Motivation Scale, as seen in Table 12. While there is no non-invariant parameter in the factor parameters (0.0%), 11 parameters are non-invariant according to the item intercepts (8.7%). Since the percentage of non-invariant parameters is lower than 25%, it can be stated that the scale provides approximate measurement invariance (Asparouhov & Muthén, 2014).

Table 13 presents the invariance findings of intercept and factor loading parameters on an item and country basis. Countries in brackets and bold are labelled as non-invariant since they show more variability than the tolerance parameter.

Table 13

Each Item's Alignment Results of 14 Countries

Item	Intercepts	Factor Loadings
BSBM16A	36, 152, 348, 380, 392, 504, 620, 643, 682, 702, 710, 752, 792, 926	36, 152, 348, 380, 392, 504, 620, 643, 682, 702, 710, 752, 792, 926
BSBM16B	36, 152, 348, 380, 392, **(504)**, 620, 643, 682, 702, 710, 752, **(792)**, 926	36, 152, 348, 380, 392, 504, 620, 643, 682, 702, 710, 752, 792, 926
BSBM16C	36, 152, 348, 380, 392, 504, **(620)**, 643, **(682)**, 702, 710, **(752)**, **(792)**, 926	36, 152, 348, 380, 392, 504, 620, 643, 682, 702, 710, 752, 792, 926
BSBM16D	36, 152, 348, 380, 392, 504, **(620)**, 643, 682, 702, 710, 752, 792, 926	36, 152, 348, 380, 392, 504, 620, 643, 682, 702, 710, 752, 792, 926
BSBM16E	36, 152, 348, 380, 392, 504, 620, 643, 682, 702, 710, 752, 792, 926	36, 152, 348, 380, 392, 504, 620, 643, 682, 702, 710, 752, 792, 926
BSBM16F	36, 152, 348, 380, 392, 504, 620, 643, 682, 702, **(710)**, 752, 792, **(926)**	36, 152, 348, 380, 392, 504, 620, 643, 682, 702, 710, 752, 792, 926
BSBM16G	36, 152, 348, 380, 392, 504, 620, 643, 682, 702, 710, 752, 792, 926	36, 152, 348, 380, 392, 504, 620, 643, 682, 702, 710, 752, 792, 926
BSBM16H	36, 152, 348, 380, 392, **(504)**, 620, 643, 682, 702, 710, 752, **(792)**, 926	36, 152, 348, 380, 392, 504, 620, 643, 682, 702, 710, 752, 792, 926
BSBM16I	36, 152, 348, 380, 392, 504, 620, 643, 682, 702, 710, 752, 792, 926	36, 152, 348, 380, 392, 504, 620, 643, 682, 702, 710, 752, 792, 926

Note. 36: Australia, 152: Chile, 348: Hungary, 380: Italy, 392: Japan, 504: Morocco, 620: Portugal, 643: Russia, 682: Saudi Arabia, 702: Singapore, 710: South Africa, 752: Sweden, 792: Turkey, 926: Pakistan.

When Table 13 is examined, it is seen that all factor loadings of Mathematics Intrinsic Motivation are invariant in 14 countries at the 8th-grade level. According to the intercept coefficients, two parameters in item BSBM16B, four parameters in item BSBM16C, one parameter in item BSBM16D, two parameters in item BSBM16F, and one parameter in item BSBM16H are non-invariant. Findings on an item basis show that 4 out of 9 items are invariant across all countries, whereas in 5 items, there are

differences in the intercept coefficients of some countries. On the other hand, it can be said that the negatively rooted items BSBM16B and BSBM16C are the items that violate the invariance the most. The country-based findings suggest that the factor loading parameters are invariant across all countries. According to the intercept parameter, Morocco is non-invariant in two items: Portugal in two, Saudi Arabia in one, South Africa in one, Sweden in one, Turkey in three, and Pakistan in one item. Both factor loadings and intercept parameters are equivalent for Australia, Chile, Hungary, Italy, Japan, Russia, and Singapore. As Asparouhov and Muthén (2014) suggested, since the non-invariant parameter ratio of factor loading and intercept parameters is below 25%, it can be stated that the Mathematics Intrinsic Motivation scale provides approximate invariance at the 8th-grade level, and all groups can be compared.

Discussion and Conclusion

The fact that the factor structures of the groups are different from each other in the measurement made with the same scale on two or more groups may indicate that the measurement tool is perceived differently in each group or that it measures a different psychological characteristic. If this is the case, it would not be correct to compare the measurements obtained from these groups. The reason for the different factor structures of the groups may be application conditions, cultural reasons, and linguistic translation problems (Başusta, 2010). According to the results of measurement invariance, it can be evaluated whether the items in the measurement tool are perceived in the same way by the participants according to groups such as different genders, languages, and cultures.

In this study, it was examined whether the mathematics intrinsic motivation scale in the TIMSS 2019 application showed measurement invariance according to 14 different countries. The analyses were performed with MG-CFA and MG Alignment methods. Since the stages in measurement invariance with MG-CFA are nested models, $\Delta\chi^2$ and ΔCFI values are generally used in the literature to compare the models. In this study, ΔCFI value was used to decide the measurement invariance between the stages due to the large sample size. Other goodness of fit statistics is also reported.

According to the measurement invariance results of the study with MG-CFA, the TIMSS 2019 4th grade mathematics intrinsic motivation scale provides only configural invariance across 14 different countries. It was observed that metric, scalar, and strict invariance were not achieved. Similarly, the 8th-grade mathematics intrinsic motivation scale provides configural and metric invariance across 14 different countries. However, it was observed that it did not provide scalar and strict invariance. Following these results, measurement invariance was analysed with the Alignment method. The results of the alignment method applied to factor loading and intercept parameters showed that the mathematics intrinsic motivation scale at the 4th-grade level provided approximate measurement invariance and that the scores of the countries could be compared. Similarly, at the 8th-grade level, the results showed that the mathematics intrinsic motivation scale also provided approximate measurement invariance and that the scores of the countries could be compared.

In the literature, there are studies examining the measurement invariance of different scales used in large-scale assessments across countries. He et al. (2019) examined the measurement invariance of 29 countries according to Instrumental Motivation, Enjoyment of Science, and Sense of School Belonging scales with PISA

and TIMSS data. The results show that all three scales provide only metric invariance. Raižienė et al. (2021) examined the measurement invariance of the motivational constructs (competitiveness, work mastery, and fairness of failure) model with the PISA 2018 application with EU countries and concluded that the model provides metric invariance. Tekin and Çobanoğlu-Aktan (2021) conducted a measurement invariance study comparing Turkey, Norway, and Singapore on the collaborative problem-solving scale with PISA 2015 data. The results of the study reported that these three countries met only the configural invariance stage. Ersözlü et al. (2022) reported that metric invariance and partial scalar invariance of the mathematics anxiety scale were met in their study examining measurement invariance between Australian and Russian groups. Karakoç-Alatlı et al. (2016) examined the measurement invariance of the TIMSS 2011 mathematics test between four different countries (Turkey, England, Japan, and the USA) and concluded that the test provided only configural invariance. According to measurement invariance with MG-CFA, the results of the current study are similar to the results of these studies. However, it should be noted that these studies did not examine measurement invariance with the Alignment method but only with MG-CFA. On the other hand, MG-CFA is not practical in comparing a large number of groups (i.e., countries) since it makes many pairwise comparisons between countries and may produce incorrect results (Sırgancı et al., 2020). In addition, when measurement invariance cannot be achieved with MG-CFA, it does not provide any information about the groups that violate invariance. Therefore, measurement invariance should be examined with the Alignment method, which overcomes these disadvantages of MG-CFA (Asparouhov & Muthén, 2014).

Glasgow et al. (2019) examined the measurement invariance of mathematics teachers' Job satisfaction, School emphasis on academic success, School condition and resources, Safe and orderly school, and teacher Self-efficacy scales obtained from TIMMS 2015 teacher surveys across 46 countries. The results showed that only three constructs achieved metric invariance, while the other constructs remained in configural invariance. In the present study, the results of measurement invariance using the alignment method show that all five constructs achieve approximate invariance, indicating that these constructs can be validly compared across educational systems. Sirganci et al. (2020), in their study explaining the basic concepts and processes of the alignment method, examined the measurement invariance of 56 countries on the Instrumental Motivation Scale data in the PISA 2015 cycle, MG-CFA results show that the scale provides only configural invariance. The measurement invariance findings with the alignment method provide more detailed information about which countries and which items contribute better to measurement invariance. Kaya et al. (2023) examined the measurement invariance of the Self-Efficacy scale data obtained from PISA 2018 on 79 countries. The researchers report that only configural invariance was met in the cross-country comparison. Jami and Kemmelmeier (2021), in their study examining the measurement invariance of the Subjective well-being scale in 36 countries, MG-CFA findings showed that metric invariance was achieved, while the alignment method was found to provide approximate measurement invariance of the scale. The results of the current study are similar to the results of these studies in the literature (Glasgow et al., 2019; Sırgancı et al., 2020; Jami & Kemmelmeier, 2020). Ertürk and Oyar (2021) examined the measurement invariance between different

countries with different methods using the "Mathematics Liking Scale" obtained from PISA 2015 data. Researchers stated that when determining the method to be used in measurement invariance studies, attention should be focused on meeting the assumptions and the structure of the data (Ertürk & Oyar, 2021). Therefore, in line with both the findings of this study and Sırgancı et al. (2020), the Alignment Method should be preferred in examining the measurement invariance of scales/tests with a large number of groups.

When the measurement invariance results on an item basis were analysed by alignment method, it was concluded that four items at the 4th-grade level and four items at the 8th-grade level were invariant according to both factor loadings and intercept parameters. On the other hand, the second and third items in the scale have negatively worded items, and it can be seen that the intercept parameters of these two items are the most non-invariant parameters according to the intercept parameter among the countries. This result shows that items negatively worded have a decreasing effect on measurement invariance. Therefore, it is important to translate the items with negative roots in a more comprehensible way and in accordance with the student's level.

When the results of measurement invariance by alignment method on a country basis were analysed, it was concluded that the parameters of 14 countries were equivalent in six countries at the 4th-grade level and seven countries at the 8th-grade level according to both factor loadings and intercept parameters. Morocco, Portugal, Saudi Arabia, and South Africa were found to have non-invariant parameters in the intercept parameters at both grade levels.

In conclusion, as Asparouhov and Muthén (2014) suggested, since the factor loading and the non-invariant parameter ratio of the intercept parameters are below 25%, the Mathematics Intrinsic Motivation scale provides approximate measurement invariance at both 4th and 8th-grade levels and the scores of the countries are comparable.

Limitations and Recommendations

In this study, the measurement invariance of the mathematics intrinsic motivation scale was compared across countries at both the 4th and 8th grades. Therefore, countries that completed this scale at both 4th and 8th-grade levels were selected, and other countries were not included in the scope of the study. This is a limitation of the study. Future studies can be carried out to cover all countries according to one grade level of this scale.

Scores obtained from large-scale exams such as PISA and TIMSS can be compared with the scores obtained by applying the same scale in different grades with measurement invariance, and the differentiation of the factor structure of different grade levels can be examined.

In this study, it was concluded that questions with negative roots had a disruptive effect on measurement invariance. The effect of negatively rooted questions on measurement invariance can be addressed in future studies.

The findings of this study - similar to the literature - show that in cross-country comparisons, MG-CFA findings do not indicate the presence of measurement invariance, while the findings of the alignment method indicate approximate

measurement invariance. Therefore, it is recommended that measurement invariance be tested with the alignment method in future multi-group measurement invariance studies.

Conflicts of Interest

There is no conflict of interest in the research.

Author Bio:

Mahmut Sami Yiğiter completed his master's degree at Ankara Hacı Bayram Veli University. He continues his doctoral studies in Hacettepe University in the field of Measurement and Evaluation in Education. He is currently working as a lecturer at Social Sciences University of Ankara. His research focuses on measurement and evaluation in education, measurement invariance, cross-cultural comparisons, meta-analysis and methodological methods.

References

Adıbatmaz, F. B. K., & Yildiz, H. (2020). The effects of distractors to differential item functioning in peabody picture vocabulary test. *Journal of Theoretical Educational Science*, *13*(3), 530-547.

Afanador, N. L., Tran, T., Blanchet, L., & Baumgartner, R. (2016). mvdalab-package 3.

Ahmed, W., Minnaert, A., Van der Werf, G., & Kuyper, H. (2010). Perceived social support and early adolescents' achievement: The mediational roles of motivational beliefs and emotions. *Journal of Youth and Adolescence*, *39*(1), 36–46. doi:10.1007/s10964-008-9367-7

Akben-Selcuk, E. (2017). Personality, motivation, and math achievement among Turkish students: Evidence from PISA data. *Perceptual and Motor Skills*, *124*(2), 514–530. https://doi.org/10.1177/0031512516686505

Arikan, S., Özer, F., Şeker, V., & Ertaş, G. (2020). The importance of sample weights and plausible values in large-scale assessments. *Journal of Measurement and Evaluation in Education and Psychology*, *11*(1), 43-60. doi: https://doi.org/10.21031/epod.602765

Barak, M., & Asad, K. (2012). Teaching image-processing concepts in junior high schools: Boys' and girls' achievements and attitudes towards technology. *Research in Science and Technological Education, 30*(1), 81–105. https://doi.org/10.1080/02635143. 2012. 656084

Başusta, N. B., & Gelbal, S. (2015). Examination of measurement invariance at groups' comparisons: A study on PISA student questionnaire. *Hacettepe University Education Faculty Journal*, *30*(4), 80-90.

Büyüköztürk, Ş., Çakmak, E. K., Akgün, Ö. E., Karadeniz, Ş., & Demirel, F. (2017). *Scientific research methods*. Pegem.

Cardoso, M. E. (2020). Policy evidence by design: International large-scale assessments and grade repetition. *Comparative Education Review*, *64*(4), 598-618. https://doi.org/10.1086/710777

Cheung G. W., Rensvold R. B. (1999). Testing factorial invariance across groups: A reconceptualization and proposed new method. *Journal of Management*, *25*(1), 1-27. https://doi.org/10.1177/014920639902500101

Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling*, *9*(2), 233-255.

Cleary, T. J., & Chen, P. P. (2009). Self-regulation, motivation, and math achievement in middle school: Variations across grade level and math context. *Journal of School Psychology*, *47*(5), 291–314. https://doi.org/10.1016/j.jsp.2009.04.002

Engel, L. C., & Rutkowski, D. (2021). Costs of big data. In C. Wyatt-Smith, B. Lingard, & E. Heck (Eds.), *Digital disruption in teaching and testing* (pp. 124–135). Routledge.

Ertürk, Z., & Erdinç-Akan, O. (2018). TIMSS 2015 matematik başarısı ile ilgili bazı değişkenlerin cinsiyete göre ölçme değişmezliğinin incelenmesi. *Journal of Theoretical Educational Science, UBEK-2018,* 204-226.

Fischer, R., & Karl, J. A. (2019). A primer to (cross-cultural) multi-group invariance testing possibilities in R. *Frontiers in Psychology*, *10*, 1507. https://doi.org/10.3389/fpsyg.2019.01507

Forero, C. G., Maydeu-Olivares, A., & Gallardo-Pujol, D. (2009). Factor analysis with ordinal indicators: A Monte Carlo study comparing DWLS and ULS estimation. *Structural Equation Modeling: A Multidisciplinary Journal*, *16*(4), 625–641. https://doi.org/10.1080/10705510903203573

Freiberger, V., Steinmayr, R., & Spinath, B. (2012). Competence beliefs and perceived ability evaluations: How do they contribute to intrinsic motivation and achievement?. *Learning and Individual Differences*, *22*(4), 518-522. https://doi.org/10.1016/j.lindif.2012.02.004

Glassow, L. N., Rolfe, V., & Hansen, K. Y. (2021). Assessing the comparability of teacher-related constructs in TIMSS 2015 across 46 education systems: an alignment optimization approach. *Educational Assessment Evaluation and Accountability*, *33*(1), 105–137. https://doi.org/10.1007/s11092-020-09348-2

Gregorich, S. E. (2006). Do self-report instruments allow meaningful comparisons across diverse population groups? Testing measurement invariance using the confirmatory factor analysis framework. *Medical Care*, *44*(11 Suppl 3), S78.

Guo, J., Parker, P. D., Marsh, H. W., & Morin, A. J. S. (2015). Achievement, motivation, and educational choices: A longitudinal study of expectancy and value using a multiplicative perspective. *Developmental Psychology*, *51*(8), 1163–1176. https://doi.org/10.1037/a0039440

Gustafsson, J.-E. (2018). International large scale assessments: Current status and ways forward. *Scandinavian Journal of Educational Research*, *62*(3), 328–332. https://doi.org/10.1080/00313831.2018.1443573

He, J., Barrera-Pedemonte, F., & Buchholz, J. (2019). Cross-cultural comparability of noncognitive constructs in TIMSS and PISA. *Assessment in Education Principles Policy and Practice*, *26*(4), 369–385. https://doi.org/10.1080/0969594x.2018.1469467

Henderson, R. W. & Landesman, E. M. (1995). Effects of thematically integrated mathematics instruction on students of Mexican descent. *Journal of Educational Research*, *88*(5), 290–300.

Hooper, M., Mullis, I. V., Martin, M. O., & Fishbein, B. (2020). TIMSS 2019 context questionnaire framework. *TIMSS*, 59-78.

Horn, J. L., & McArdle, J. J. (1992). A practical and theoretical guide to measurement invariance in aging research. *Experimental Aging Research*, *18*(3–4), 117–144. https://doi.org/10.1080/03610739208253916

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

Ilter, İ. (2021). The relationship between academic amotivation and academic achievement: A study on middle school students. *Kuramsal Eğitimbilim Dergisi*, *14*(3), 389–410. https://doi.org/10.30831/akukeg.847145

İlhan, M., & Çetin, B. (2013). Matematik odaklı epistemolojik inanç ölçeği (MOEİÖ): Geçerlik ve güvenirlik çalışması. *Kurumsal Eğitimbilim Dergisi*, *6*(3), 362-368.

Jami, W. A., & Kemmelmeier, M. (2020). Assessing well-being across space and time: Measurement equivalence of the WHO-5 in 36 European countries and over 8 years. *Journal of Well-Being Assessment*, *4*(3), 419–445. https://doi.org/10.1007/s41543-021-00042-8

Jöreskog, K. G., & Sörbom, D. (1993). *LISREL 8: Structural equation modeling with the SIMPLIS command language*. Scientific software international.

Kaliyaperumal, S. K., Kuppusamy, M., & Gounder, A. S. (2015). Outlier detection and missing value in time series ozone data. *International Journal of Scientific Research in Knowledge*, *3*(9), 220–226. https://doi.org/10.12983/ijsrk-2015-p0220-0226

Kam, C. C. S. (2019). Careless responding threatens factorial analytic results and construct validity of personality measure. *Frontiers in Psychology*, *10*, 1258. https://doi.org/10.3389/fpsyg.2019.01258

Kam, C. C. S., & Meyer, J. P. (2015). How careless responding and acquiescence response bias can influence construct dimensionality: The case of job satisfaction. *Organizational Research Methods*, *18*(3), 512–541. https://doi.org/10.1177/1094428115571894

Kaya, S., Eryilmaz, N., & Yuksel, D. (2023). A cross-cultural comparison of self-efficacy as a resilience measure: Evidence from PISA 2018. *Youth & Society*. https://doi.org/10.1177/0044118x231186833

Kline, R. B. (2011). Convergence of structural equation modeling and multilevel modeling. In W. Vogt & M. Williams (Eds.), The SAGE handbook of innovation in social research methods (pp. 562–589). SAGE Publications. https://doi.org/10.4135/9781446268261.

Koğar, H., & Yilmaz Koğar, E. (2015). Comparison of different estimation methods for categorical and ordinal data in confirmatory factor analysis. *Journal of Measurement and Evaluation in Education and Psychology*, *6*(2), 351-364. https://doi.org/10.21031/epod.94857

Malone, T. W., & Lepper, M. R. (2021). Making learning fun: A taxonomy of intrinsic motivations for learning. In R. E. Snow & M. J. Farr (Eds.), *Aptitude, learning, and instruction* (pp. 223-254). Routledge.

Meredith, W. (1993). Measurement invariance, factor analysis and factorial invariance. *Psychometrika*, *58*(4), 525–543.

Middleton, J. (1995). A study of intrinsic motivation in the mathematics classroom: A personal constructs approach. *Journal for Research in Mathematics Education, 26*(3), 254–279. https://doi.org/10.2307/749130

Milfont, T. L., & Fischer, R. (2010). Testing measurement invariance across groups: Applications in cross-cultural research. *International Journal of Psychological Research*, *3*(1), 111-130.

Millsap, R. E., & Olivera-Aguilar, M. (2012). Investigating measurement invariance using confirmatory factor analysis. In R. H. Hoyle, (Ed.), *Handbook of structural equation modeling* (pp. 380-392). Guilford.

Mueller, M., Yankelewitz, D., & Maher, C. (2011). Sense making as motivation in doing mathematics: Results from two studies. *The Mathematics Educator*, *20*(2), 33–43.

Mullis, I. V. S., & Martin, M. O. (Eds.). (2017). *TIMSS 2019 Assessment Frameworks*. Retrieved from Boston College, TIMSS and PIRLS International Study Center website: http://timssandpirls.bc.edu/timss2019/frameworks/

Muthén ©n, B., & Asparouhov, T. (2014). IRT studies of many groups: The alignment method. *Frontiers in Psychology*, *5*. https://doi.org/10.3389/fpsyg.2014.00978

Nugent, G., Barker, B., Grandgenett, N., & Adamchuk, V. (2010). Impact of robotics and geospatial technology interventions on youth stem learning and attitudes. *Journal of Research on Technology in Education, 42*(4), 391–408.

OECD. (2013). *Students' drive and motivation.* In PISA 2012 results: Ready to learn (Volume III): Students' engagement, drive, and self-beliefs. OECD Publishing. Retrieved from http://dx.doi.org/10.1787/ 9789264201170-7-en.

Peterson, B. G., Carl, P., Boudt, K., Bennett, R., Ulrich, J., Zivot, E., ... & Wuertz, D. (2018). Package 'performanceanalytics'. *R Team Cooperation*, *3*, 13-14.

Putnick, D. L., & Bornstein, M. H. (2016). Measurement invariance conventions and reporting: The state of the art and future directions for psychological research. *Developmental Review: DR*, *41*, 71–90. https://doi.org/10.1016/j.dr.2016.06.004

Raižienė, S., Ringienė, L., Laukaityte, I., & Jakaitienė, A. (2021). Measurement invariance of pisa 2018 motivational constructs across EU countries. *EDULEARN21 Proceedings* (pp. 7081-7081). IATED.

Raykov, T. (2004). Behavioral scale reliability and measurement invariance evaluation using latent variable modeling. *Behavior Therapy*, *35*(2), 299–331. https://doi.org/10.1016/s0005-7894(04)80041-8

Robitzsch, A. (2020). sirt: Supplementary item response theory models. R package version 3.4-64. https://CRAN.R-project.org/package=sirt

Rosseel, Y. (2012). lavaan: An R Package for Structural Equation Modeling. *Journal of Statistical Software*, *48*(2). https://doi.org/10.18637/jss.v048.i02

Rosseel, Y. (2012). lavaan: An*R*Package for Structural Equation Modeling. *Journal of Statistical Software*, *48*(2). https://doi.org/10.18637/jss.v048.i02

Rutkowski, L., & Svetina, D. (2014). Assessing the hypothesis of measurement invariance in the context of large-scale international surveys. *Educational and Psychological Measurement*, *74*(1), 31–57. https://doi.org/10.1177/0013164413498257

Ryan, R.M., & Deci E.L., (2009) Promoting self-determined school engagement: motivation, learning, and well-being. In D. B. Miele & K. R. Wentzel (Eds.), *Handbook on motivation at school* (pp 171–196). Routledge.

Schmitt, N., & Kuljanin, G. (2008). Measurement invariance: Review of practice and implications. *Human Resource Management Review, 18*(4), 210–222. https://doi.org/10.1016/j.hrmr.2008.03.003

Shores, M. L., & Shannon, D. M. (2007). The effects of self-regulation, motivation, anxiety, and attributions on mathematics achievement for fifth and sixth grade students. *School Science and Mathematics*, *107*(6), 225–236. Retrieved from http://ssmj.tamu.edu

Sırgancı, G., Uyumaz, G., & Yandi, A. (2020). Measurement invariance testing with alignment method: Many groups comparison. *International Journal of Assessment Tools in Education, 7*(4), 657–673. https://doi.org/10.21449/ijate.714218

Sözer, E., Eren, B., & Kahraman, N. (2021). Investigating measurement invariance for longitudinal assessments: An application using repeated data over four weeks. *Gazi Üniversitesi Gazi Eğitim Fakültesi Dergisi*, *41*(2), 729–763. https://doi.org/10.17152/gefad.873885

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

Taris, T. W., Bok, I. A., & Meijer, Z. Y. (1998). Assessing stability and change of psychometric properties of multi-item concepts across different situations: A general approach. *The Journal of Psychology*, *132*(3), 301–316. https://doi.org/10.1080/00223989809599169

Tavani C.M., & Losh S.C. (2003) Motivation, self-confidence, and expectations as predictors of the academic performances among our high school students. *Child Study J, 33*(3), 141–151.

Tierney, N., Cook, D., McBain, M., & Fay, C. (2021). naniar: Data structures, summaries, and visualisations for missing data (R package version 0.6.1)[Computer software].

Uyar, Ş. & Doğan, N. (2014). An investigation of measurement invariance of learning strategies model across different groups in PISA Turkey sample. *International Journal of Turkish Education Sciences, 2014*(3), 30-43.

Van De Schoot, R., Schmidt, P., De Beuckelaer, A., Lek, K., & Zondervan-Zwijnenburg, & M. (2015). Editorial: Measurement invariance. *Frontiers in Psychology, 6*, 1064. https://doi.org/10.3389/fpsyg.2015.01064

Vandenberg, R. J., & Lance, C. E. (2000). A review and synthesis of the measurement invariance literature: Suggestions, practices, and recommendations for organizational research. *Organizational Research Methods*, *3*(1), 4–70. https://doi.org/10.1177/109442810031002

Weidinger, A. F., Steinmayr, R., & Spinath, B. (2017). Math grades and intrinsic motivation in elementary school: A longitudinal investigation of their association. *British Journal of Educational Psychology*, *87*(2), 187-204.

Wickham, H., François, R., Henry, L., Müller, K., & Wickham, M. H. (2019). Package 'dplyr'. *A Grammar of Data Manipulation. R package version*, *8*.

Woods, C. M. (2006). Careless responding to reverse-worded items: Implications for confirmatory factor analysis. *Journal of Psychopathology and Behavioral Assessment*, *28*(3), 186-191.

Woolley, M. E., Strutchens, M. E., Gilbert, M. C., & Martin, W. (2010). Mathematics success of black middle school students: Direct and indirect effects of teacher expectations and reform practices. *Negro Educational Review*, *61*(1), 41–59.

Wu, A. D., Li, Z., & Zumbo, B. D. (2007). *Decoding the meaning of factorial invariance and updating the practice of multi-group confirmatory factor analysis: A demonstration with TIMSS data*. University of Massachusetts Amherst. https://doi.org/10.7275/MHQA-CD89

Yiğiter, M. S. (2023). Matematik duyuşsal özellik faktörlerinin cinsiyete göre ölçme değişmezliğinin incelenmesi: TIMSS 2019 Türkiye örneği. *Anadolu Üniversitesi Eğitim Fakültesi Dergis*i, *7*(4), 859–882. https://doi.org/10.34056/aujef.1198134

Yildirim, S. (2011). Self-efficacy, intrinsic motivation, anxiety and mathematics achievement: Findings from Turkey, Japan and Finland. *Necatibey Faculty of Education Electronic Journal of Science and Mathematics Education*, *5*(1), 277-291.

Yin, L., & Fishbein, B. (2019). Creating and interpreting the TIMSS 2019 context questionnaire scales. *Methods and procedures: TIMSS*, 16-1.

Zembat, R., Akşin-Yavuz, E., Tunçeli, H. İ., Yılmaz, H. (2018). Öğretmenlik mesleğine yönelik tutum ile akademik motivasyon ve başarı arasındaki ilişkinin incelenmesi. *Kuramsal Eğitimbilim Dergisi [Journal of Theoretical Educational Science], 11*(4), 789-808.

Appendix 1

CFA Result (4th Grade)

Country	i1	i2	i3	i4	i5	i6	i7	i8	i9	cfi	tli	rmsea	srmr	Cronbach	Omega	AVE
Australia	0.826	0.750	0.824	0.565	0.930	0.732	0.838	0.933	1.007	0.998	0.998	0.030	0.025	0.941	0.944	0.657
Chile	0.680	0.547	0.661	0.473	0.851	0.627	0.773	0.869	0.912	0.985	0.980	0.063	0.053	0.899	0.903	0.519
Hungary	0.849	0.647	0.675	0.520	0.936	0.721	0.720	0.938	0.970	0.995	0.994	0.042	0.036	0.919	0.923	0.581
Italy	0.802	0.760	0.757	0.442	0.864	0.697	0.681	0.895	0.932	0.998	0.997	0.026	0.025	0.926	0.930	0.605
Japan	0.744	0.417	0.531	0.602	0.899	0.711	0.803	0.816	0.865	0.996	0.995	0.044	0.044	0.934	0.939	0.642
Morocco	0.331	0.262	0.297	0.403	0.471	0.408	0.470	0.516	0.520	0.921	0.894	0.062	0.067	0.778	0.768	0.278
Portugal	0.667	0.583	0.684	0.458	0.811	0.609	0.735	0.787	0.881	0.993	0.991	0.041	0.043	0.919	0.922	0.577
Russia	0.686	0.492	0.606	0.446	0.741	0.642	0.663	0.798	0.870	0.997	0.996	0.033	0.032	0.905	0.909	0.536
Saudi Arabia	0.547	0.359	0.487	0.555	0.706	0.576	0.667	0.746	0.751	0.954	0.938	0.072	0.066	0.863	0.862	0.420
Singapore	0.740	0.723	0.724	0.574	0.858	0.671	0.810	0.809	0.945	0.998	0.997	0.033	0.030	0.933	0.936	0.621
South Africa	0.517	0.285	0.340	0.508	0.656	0.603	0.593	0.622	0.667	0.934	0.911	0.080	0.062	0.796	0.794	0.312
Sweden	0.722	0.699	0.881	0.655	0.938	0.756	0.775	0.898	0.974	0.998	0.997	0.033	0.029	0.944	0.946	0.664
Turkey	0.540	0.526	0.633	0.325	0.657	0.449	0.634	0.735	0.780	0.988	0.984	0.038	0.045	0.885	0.889	0.484
Pakistan	0.768	0.680	0.768	0.529	0.879	0.719	0.800	0.909	0.977	0.997	0.996	0.034	0.030	0.935	0.938	0.633

Note. i1: item1 factor loading, AVE : average variance extracted

Appendix 2

CFA Result (8th Grade)

Country	i1	i2	i3	i4	i5	i6	i7	i8	i9	cfi	tli	rmsea	srmr	Cronbach	Omega	AVE
Australia	0.863	0.722	0.692	0.636	0.926	0.724	0.843	0.795	0.944	0.997	0.996	0.041	0.031	0.938	0.940	0.641
Chile	0.748	0.610	0.696	0.558	0.900	0.701	0.829	0.753	0.904	0.994	0.992	0.054	0.045	0.930	0.932	0.609
Hungary	0.875	0.635	0.671	0.647	0.968	0.805	0.802	0.734	0.852	0.996	0.995	0.045	0.035	0.932	0.934	0.614
Italy	0.929	0.808	0.815	0.587	1.039	0.824	0.869	0.790	1.056	0.999	0.999	0.032	0.023	0.948	0.951	0.688
Japan	0.795	0.515	0.582	0.669	0.902	0.666	0.812	0.697	0.866	0.995	0.994	0.048	0.043	0.942	0.945	0.664
Morocco	0.733	0.486	0.598	0.668	0.862	0.682	0.722	0.857	0.903	0.990	0.986	0.056	0.048	0.908	0.911	0.540
Portugal	0.869	0.774	0.776	0.694	0.982	0.724	0.892	0.836	1.004	0.998	0.997	0.035	0.027	0.947	0.948	0.674
Russia	0.694	0.518	0.568	0.545	0.792	0.656	0.733	0.675	0.844	0.996	0.994	0.043	0.038	0.922	0.924	0.582
Saudi Arabia	0.790	0.557	0.711	0.750	0.994	0.832	0.903	0.936	1.035	0.991	0.989	0.068	0.053	0.928	0.931	0.607
Singapore	0.767	0.754	0.663	0.554	0.862	0.655	0.818	0.724	0.984	0.997	0.996	0.041	0.034	0.935	0.938	0.632
South Africa	0.632	0.546	0.563	0.514	0.812	0.651	0.758	0.681	0.865	0.988	0.984	0.054	0.045	0.890	0.893	0.488
Sweden	0.809	0.719	0.801	0.669	0.925	0.798	0.828	0.744	0.910	0.996	0.995	0.046	0.035	0.943	0.945	0.657
Turkey	0.700	0.721	0.809	0.519	0.897	0.752	0.840	0.872	1.016	0.993	0.991	0.053	0.041	0.918	0.921	0.571
Pakistan	0.777	0.703	0.687	0.611	0.868	0.706	0.824	0.768	0.887	0.997	0.996	0.043	0.035	0.933	0.934	0.614

Note. i1: item1 factor loading, AVE : average variance extracted

Journal of Theoretical Educational Science, 17(1), 28-56, January 2024
Kuramsal Eğitimbilim Dergisi, 17(1), 28-56, Ocak 2024
[Online]: http://dergipark.org.tr/akukeg
DOI number: http://doi.org/10.30831/akukeg.1295046

Trends, Opportunities, and Pipelining in Preschool STEM Education: A Scoping Review

Okul Öncesi STEM Eğitiminde Eğilimler, Fırsatlar ve Eksiklikler: Bir Kapsam Belirleme İncelemesi

İsmail DÖNMEZ* **Salih GÜLEN****

Received: 10 May 2023 **Research Article** **Accepted:** 27 November 2023

ABSTRACT: This study was planned with the aim of conducting a comprehensive literature analysis of preschool STEM education research and evaluating future opportunities. Scoping Review was conducted by analyzing 29 articles in journals published by Springer, Taylor & Francis, Elsevier, and SAGE, which were obtained using combinations of the keywords "STEM," "STEM education," "kindergarten," and "preschool." The results show that the literature on STEM research in preschool, which has largely developed in the United States, has grown in recent years, forming a relatively new and expanding field. It is concluded that preschool STEM education research promises to be a popular field in the future. Studies in which different STEM activities were carried out were identified along with the examinations. In these studies, activities focus on games, teacher roles, or program and model development. For such activities, factors such as the role of sample groups in the process and their impact on the process should be determined. In other words, there is a need for studies focusing on the evaluation dimension of STEM education in preschool education.

Keywords: Preschool, scoping review, STEM, STEM education.

ÖZ: Bu çalışma, okul öncesi STEM eğitimi araştırmalarına ilişkin kapsamlı bir literatür analizi yapmak ve gelecekteki fırsatları değerlendirmek amacıyla planlanmıştır. Kapsam Belirleme İncelemesi, Springer, Taylor & Francis, Elsevier ve SAGE tarafından yayınlanan dergilerde yer alan ve "STEM", "STEM eğitimi", "anaokulu" ve "okul öncesi" anahtar kelimelerinin kombinasyonları kullanılarak elde edilen 29 makale incelenerek yapılmıştır. Sonuçlar, Amerika Birleşik Devletleri'nde büyük ölçüde gelişen okul öncesi STEM araştırmalarına ilişkin literatürün son yıllarda büyüyerek nispeten yeni ve genişleyen bir alan oluşturduğunu göstermektedir. Okul öncesi STEM eğitimi araştırmalarının gelecekte popüler bir alan olmayı vaat ettiği sonucuna varılmıştır. İncelemelerle beraber farklı STEM aktivitelerin gerçekleştiği çalışmalar tespit edilmiştir. Bu çalışmalarda aktiviteler oyunlara, öğretmen rollerine veya program ve model geliştirmeye odaklanmaktadır. Bu tür faaliyetler için örneklem gruplarının süreçteki rolü ve sürece etkisi gibi faktörlerin belirlenmesi gerekmektedir. Diğer bir deyişle, okul öncesi eğitimde STEM eğitiminin değerlendirme boyutuna odaklanan çalışmalara ihtiyaç vardır.

Anahtar kelimeler: Okul öncesi, kapsam belirleme, STEM, STEM eğitimi.

* Assoc. Prof. Dr., Muş Alparslan University, Muş, Turkey, ismaildonmezfen@gmail.com, https://orcid.org/0000-0002-7792-0169
** *Corresponding Author*: Assoc. Prof. Dr., Muş Alparslan University, Muş, Turkey, sgnova@windowslive.com, https://orcid.org/0000-0001-5092-0495

Citation Information
Dönmez, İ., & Gülen, S. (2024). Trends, opportunities, and pipelining in preschool STEM education: A scoping review. *Kuramsal Eğitimbilim Dergisi [Journal of Theoretical Educational Science], 17*(1), 28-56.

ISSN: 1308-1659

To facilitate students' learning through exploration and experimentation, the integration of science, technology, engineering, and mathematics (STEM) is an international priority (Bybee, 2013). One factor that has been identified as significantly contributing to young children's later school achievement is early exposure to STEM knowledge and skills (Morgan et al., 2016). Alongside recommendations on engaging students in STEM education, the importance of introducing children to STEM opportunities early in their education is often highlighted (National Research Council [NRC], 2011). STEM education is important in terms of nurturing and maintaining young children's natural interests in the relevant subjects and career fields and developing their problem-solving, critical, and logical thinking skills while strengthening intellectual habits (Ata-Aktürk & Demircan, 2021; Lange et al., 2019). Over the years, great importance has been placed on STEM as a field of study that is key for a country to gain a competitive advantage on the global stage, leading to a shift in educational paradigms emphasizing the importance of STEM (Kayan-Fadlelmula et al., 2022). With a simple Google search, 450 million websites can be accessed using the terms "STEM," "STEM education," and "STEM education research" (Li et al., 2020). This is because STEM education contributes to students in various ways, including academic success, attitude, and motivation. Moreover, it is recognized as an important factor for meeting future job-need expectations (Psycharis, 2018). However, academic identity development in STEM in early childhood and its potential impact on future STEM participation is largely neglected, both in research and in practice (Early Childhood STEM Working Group, 2017). A better understanding of the ways in which STEM education research is defined and related to the preschool period will contribute to the identification of trends, opportunities, and deficiencies.

Literature Review

Preschool Education

Researchers, policy-makers, and educators focus on the positive impact of high-quality early childhood education on children's development, as they do for other education levels (Brenneman et al., 2019). One of the main goals of early childhood education is to create an environment that supports lifelong learning. Therefore, one important question is what kinds of experiences in early childhood education are most valuable in the learning environment (Katz, 2010). Preschool children have a natural inclination toward science due to their sense of curiosity and ability to find solutions based on creativity and imagination (DeJarnette, 2018). During early childhood, the development of abilities such as self-regulation, working memory, and inhibitory control increases exponentially, thus establishing this period as a window of opportunity for interventions aimed at promoting child development (Tsujimoto, 2008). Furthermore, the impact of preschool education on cognitive development has been explored in studies such as Yan et al. (2021) The study found that the impact of preschool education on children's cognitive development varies depending on the cognitive ability and the length of time. This suggests that preschool education can have differential effects on different aspects of cognitive development, such as language cognition and mathematical cognition.

STEM Education

In today's modern, digitalized, and unpredictable world where knowledge changes faster than educational systems, STEM skills are perceived as the key to innovatively solving the problems of contemporary life, overcoming social and economic disparity, and achieving sustainable living (Karaşah Çakıcı et al., 2021; Jang, 2016). STEM is an interdisciplinary approach that helps students understand the concepts and contents that form the related fields based on daily-life problems. Through STEM education, students can develop 21st-century skills such as adaptability, problem-solving, communication skills, and systematic thinking (NRC, 2010). STEM education contributes to students becoming better problem-solvers, innovators, logical thinkers, inventors, and technology users (Morrison, 2006). STEM is also seen as an important field in enabling children to recognize future professions, and early STEM education shapes participation in STEM fields in the future (Campbell et al., 2020). Children have natural tendencies that enable them to learn STEM topics easily, such as the ability to make sense of experiences, analyze, hypothesize, and predict (Katz, 2010). Effective STEM education should start early, preferably from the preschool years, as it positively influences students' aspirations in relation to tertiary STEM study and STEM career pursuits (Murphy et al., 2018). However, there is a need for instructional guidelines and curricular materials for integrated STEM teaching (Guzey et al., 2016). Therefore, it is crucial to consider how to design integrated STEM activities and evaluate their effectiveness. Since there is no universal guide or model for the implementation of STEM activities, it may be difficult to implement STEM activities (Wang et al., 2011). However, field studies can serve as guides for effective practices.

Preschool STEM Education

Learning science and engineering practices in early childhood increases children's curiosity and pleasure in exploring the world around them and builds the background knowledge for science learning in the K-12 years (National Academy of Engineering and National Research Council [NSTA], 2014). One of the indicators of high-quality early childhood education is the introduction of STEM opportunities to children (NRC, 2011). Appropriate STEM experiences in early childhood may be starting points for supporting children's further success in STEM fields at the primary, secondary, and post-secondary levels (Tao, 2019). However, studies show that research on STEM education in early childhood is still in its infancy (Tippett & Milford, 2017), which is unfortunate because children are inherently naturalists and engineers (Brophy et al., 2008). Although there is an increasing awareness of the importance of early childhood STEM education and efforts are being made to include it in curricula, it is seen that teachers usually do not include interdisciplinary STEM content in their lesson plans (Tao, 2019). This is because, although teachers may know the importance of STEM education, they often do not have sufficient content knowledge and are unprepared for teaching STEM topics (Hammack & Ivey, 2017). Studies show that preschool educators rarely have in-depth knowledge of mathematics and science, and due to their insufficient content knowledge, they lack confidence in their own abilities to implement quality STEM learning experiences (Greenfield et al., 2009). A critical step to improve outcomes for children is to improve the support for educators to provide high-quality STEM experiences in the preschool period (Brenneman et al., 2019). This

can be done by creating rich content for the development of STEM activities. It has been previously demonstrated that the quality of children's learning environments before the age of six has an impact on their later academic achievements (Hadzigeorgiou, 2002). For this reason, they should learn the basic concepts and contents of STEM with hands-on activities in the preschool years within rich learning environments (Leung & Xinyun, 2019). On the other hand, effective methods for early childhood STEM education, teacher training, and curriculum design are important. One important factor is the professional preparedness of teachers in teaching STEM subjects. Research has shown that teachers' lack of self-confidence and inadequate training in STEM areas can hinder the provision of quality STEM learning experiences (Aldemir & Kermani, 2016). Therefore, providing teachers with the necessary training and support is essential to enhance their confidence and pedagogical beliefs in STEM education (Yang et al., 2021). This can be achieved through professional development programs that focus on integrating STEM practices into early childhood classrooms (Sydon & Phuntsho, 2022). In addition to teacher training, the curriculum design and learning environment also play a significant role in the effectiveness of early childhood STEM education. It is important to create developmentally appropriate STEM practices that align with the interests and abilities of young children (Çetin & Demircan, 2020). This can be achieved by incorporating play-based learning approaches, which have been effective in promoting early childhood STEM learning (Sydon & Phuntsho, 2022). Furthermore, integrating technology, such as touchscreen devices, can supplement early STEM education and enhance children's learning experiences (Aladé et al., 2016). Another aspect to consider is the development of children's STEM habits of mind. STEM process skills developed during the early years have lifelong positive effects on young children (Yang et al., 2023). Therefore, it is important to assess and foster children's STEM habits of mind through appropriate educational interventions. Moreover, career awareness and exploration can be integrated into early childhood STEM education to provide children with a broader understanding of the relevance and applications of STEM subjects (Manowaluilou & Nilsook, 2023). This can be achieved through the use of career linkage strategies that connect STEM education with real-world careers (Manowaluilou & Nilsook, 2023). Overall, effective early childhood STEM education requires a comprehensive approach that includes teacher training, curriculum design, learning environment, technology integration, and the development of children's STEM habits of mind.

Aim of the Study

It is common in educational research to conduct reviews to discover situations and trends in certain disciplines (Karampelas, 2021). In this study, research conducted on STEM education in early childhood is examined. This work will contribute to the identification of the needs and trends, opportunities, and deficiencies in STEM education in early childhood. In line with the identified needs, this study will also provide information about the contents, methods, and analytical techniques of research in the field of STEM education in early childhood, offering suggestions about research to be done in the future, as well.

The aim of this study is to synthesize the existing body of knowledge about STEM education in the preschool period and examine how the relevant factors are

related to each other. Within that framework, this research was designed with the aim of presenting the contents of educational activities based on a thorough analysis of the articles published on STEM so far. The main questions addressed in this study are as follows: Among articles published on the topic of STEM education in the preschool period, (a) what are their distributions according to publishers and publication years? (b) In which journals were they were published? (c) What are the distributions of countries and subjects? (d) Which research methods are used, and what is the typical research duration? (e) What are the distributions of sample (observation) group types and numbers? (f) What keywords are used? (g) What data collection tools are used? (h) How do the results of the articles align with the aim in the context of STEM education?

Method

It is particularly important to apply scoping review, or numerical analysis, in studies of the development and characteristics of STEM education research at more specific levels (e.g., the age group of 0-6 years) with statistical methods (Kürklü, 2019). Scoping reviews involve locating, analyzing, and presenting relevant data to gather evidence in a specific area, all done by adhering to a prescribed protocol within the existing literature (Munn et al., 2018). Although this approach is relatively new, it is used to reach conclusions by narrowing the scope of wide-ranging topics (Pham et al., 2014). This type of research enables a large amount of literature to be examined quickly and transferred to the reader. It can be used especially in answering questions made or curious about a specific subject in large time intervals (Tricco et al., 2015). The present study was designed with the aim of conducting a comprehensive literature analysis of preschool STEM education research and evaluating future opportunities. The findings provide important insight into current STEM education trends, using the current evidence base as a reference for future STEM education research and development. Scoping reviews are a type of literature review that aims to map the existing literature and provide an overview of the available evidence on a particular topic or research area (Colquhoun et al., 2014; Phan-Le et al., 2022). They are particularly useful when the topic has not been extensively reviewed, or the literature is complex and heterogeneous (Phan-Le et al., 2022). Scoping reviews differ from other types of literature reviews in terms of their purpose, methodology, and reporting. Unlike systematic reviews, which aim to answer a specific research question and provide a comprehensive synthesis of the available evidence, scoping reviews have a broader objective of mapping the literature and identifying key concepts, gaps in the research, and types and sources of evidence to inform practice, policymaking, and further research (Colquhoun et al., 2014). Scoping reviews are often conducted as standalone projects and can be undertaken when an area is complex or has not been comprehensively reviewed before (Colquhoun et al., 2014). Scoping reviews tend to focus on the breadth of existing literature rather than the depth of coverage (Arksey & O'Malley, 2005). They are an ideal tool for determining the scope or extent of literature on a particular topic and providing an overview, whether broad or detailed, as well as assessing the volume of literature and studies available (Munn et al., 2018). These reviews are especially valuable for examining emerging evidence when it is not yet clear if specific research questions can be formulated, making them a precursor to more definitive systematic reviews (Armstrong et al., 2011). The overarching goal of conducting scoping reviews is to systematically identify and

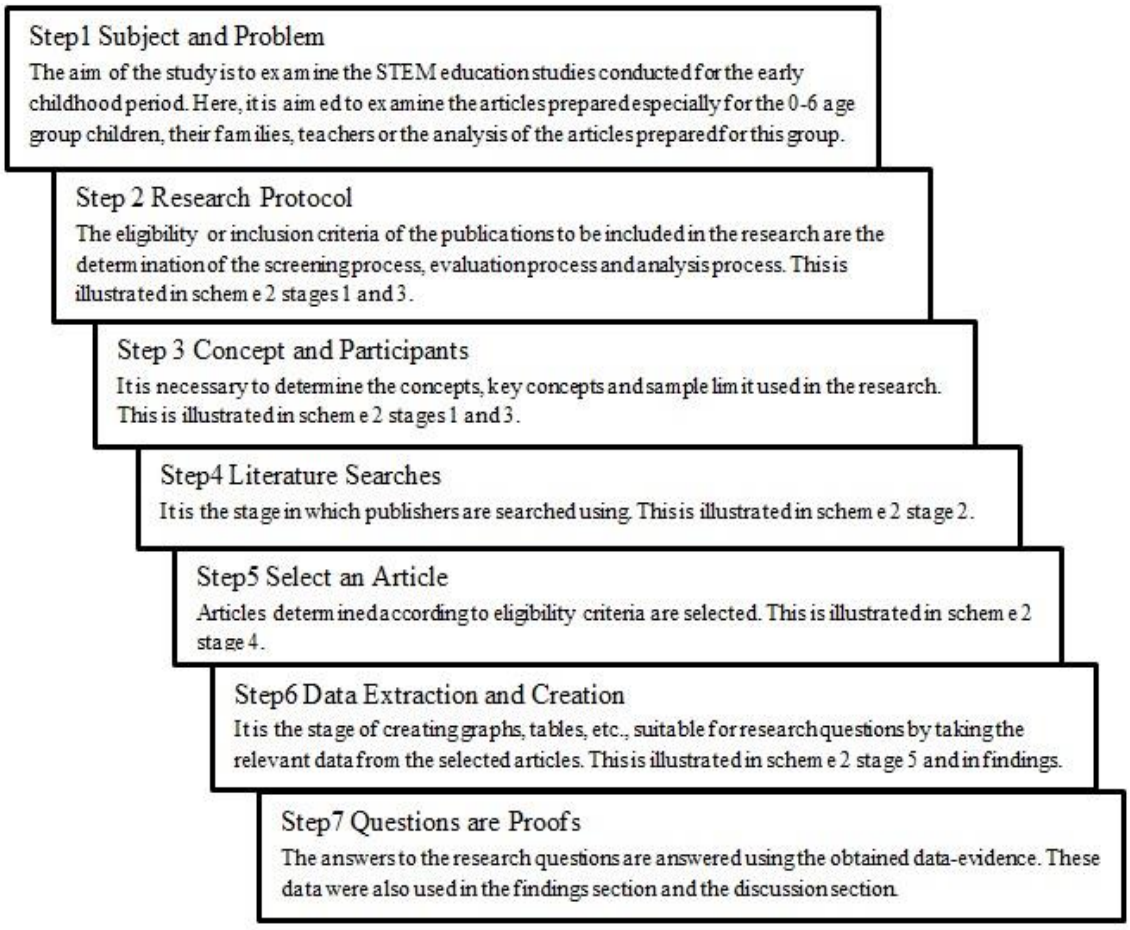
map the available evidence on a given topic, thereby providing a comprehensive landscape of the existing literature (Munn et al., 2018). While Arksey & O'Malley (2005) conducted the coverage review, we recorded the following information;

- Author(s), year of publication, place of study
- Type of intervention and comparator (if applicable); response time
- Study populations,
- Objectives of the study
- Methodology
- Outcome criteria
- Important results.

In this study, we adhered to the following procedures: "Define the aims and scope of the scoping review," "Choose the techniques for scoping review," "Collect the data," and "Run the analysis and report the findings." The steps of the plan processed in the scoping review are given below.

Schema 1

Steps Followed in Scoping Review Planning



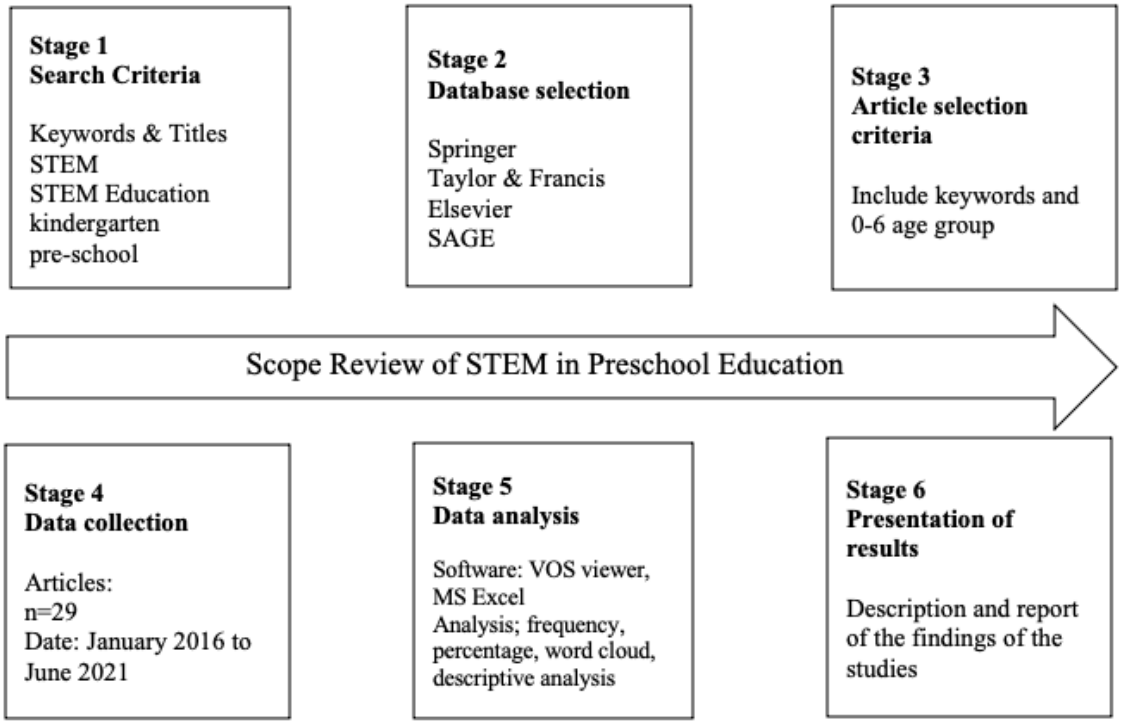
The steps followed in planning the scoping review for the research are shown in Schema 1. In this schema, seven steps of the method are followed in the research. These steps were created by researchers using Munn et al. (2018), Peters et al. (2015), and Scoping reviews (Phan-Le et al., 2022) studies.

Data Collection Process

The data collection process for this scoping review followed a systematic and rigorous approach to ensure the reliability and replicability of the study. The criteria used to select the articles to be analyzed were clearly defined to maintain the quality of both the analysis and the selected articles (Hemingway & Brereton, 2009; Moule & Goodman, 2009; Özkaya, 2018). To identify relevant articles, we focused on publications from reputable sources such as Springer, Taylor & Francis, Elsevier, and SAGE, which are known for their high-quality content in the field of STEM education. The decision to focus on these publishers aimed to target articles with significant academic rigor and relevance. Keywords such as "STEM," "STEM education," "kindergarten," and "preschool" were carefully selected based on their relevance to the research context. These keywords were then used to conduct systematic searches on the selected platforms, namely Springer, Taylor & Francis, Elsevier, and SAGE. The search process was conducted with precision and consistency to ensure that relevant articles were retrieved. To align with the research focusing on STEM education in preschool, the publication years of the selected articles were restricted to those published between January 2016 and June 2021. The choice of this timeframe was deliberate as it aimed to capture the most recent studies that would be instrumental in identifying current trends and practices. Throughout the search and selection process, adherence to the criterion that the articles analyzed should address the 0-6 age group was rigorously maintained. This criterion was central to maintaining the relevance and alignment of the study with the research objectives. As a result, 25 articles were excluded from the analysis on the grounds that they failed to meet this criterion. As a result, the scoping review focused on a comprehensive analysis of 29 articles that met all the identified criteria. A multifaceted approach was used to analyze the data. First, keywords relevant to the research scope were chosen, and selected academic publishers (Springer, Taylor & Francis, Elsevier, and SAGE) served as the primary platforms for data collection. The search process was systematically conducted on these platforms using the identified keywords. The collected publications were then subjected to a rigorous analysis based on several key criteria. These criteria included year of publication, country of origin, sample characteristics, journal of publication, topic, and relevance of the results to the research objectives. The analysis was carried out using tools such as Excel for data organization and VOS (Visualizing Output of Science) to create visual representations, concept clouds, and graphical representations to facilitate understanding of data trends and relationships. The stages followed in data collection are given in Schema 2.

Schema 2

Research Implementation Steps



Schema 2 shows the stages of the present research. The researchers completed the study within the framework of these stages.

Analysis of Data

Initially, keywords were selected in order to examine the outputs of research on preschool STEM education in detail. Springer, Taylor & Francis, Elsevier, and SAGE were then chosen for the platform. The search process was carried out on the selected platform with the determined keywords. The relevant publications were then analyzed in terms of publication year, country, samples, journal, subject, and the relationship of results and aims. Excel and the VOS program were used while analyzing the data. The obtained data are provided as frequencies, percentages, and concept clouds.

This scoping review was rigorously conducted to map the existing literature and provide an overview of the current evidence on STEM education in preschool. The research aimed to identify trends, gaps, and opportunities in this area, offering insights that can inform practice, policymaking, and future research efforts.

Findings

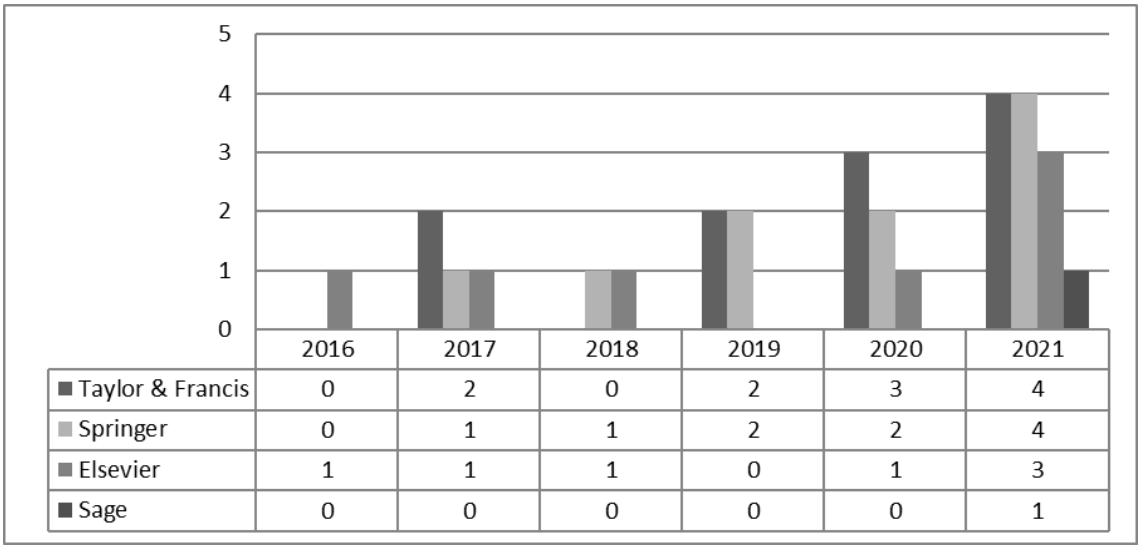
The findings obtained from this research are presented below in the order of the research questions.

Distribution of Articles by Publisher and Publication Year

The distribution of publishers for 2016-2021 is illustrated in Graph 1.

Graph 1

Number of Relevant Articles Published over the Years by Each Publisher



	2016	2017	2018	2019	2020	2021
Taylor & Francis	0	2	0	2	3	4
Springer	0	1	1	2	2	4
Elsevier	1	1	1	0	1	3
Sage	0	0	0	0	0	1

As seen in Graph 1, Springer published relevant articles in all the years under consideration, excluding 2016. Springer has the highest average number of publications per year; particularly noteworthy are its four relevant articles in 2021. Taylor & Francis did not publish any relevant articles in 2016 or 2018. However, they published a total of 11 relevant articles across the other considered years. Springer published a total of 10 relevant articles in the years under consideration. Elsevier published relevant articles in all the years considered except 2018 and 2019. SAGE only published one relevant article in 2021.

Distribution of Articles by Journals

The international journals publishing research on STEM education in early childhood are given in Table 1.

Table 1

Journals Publishing Relevant Research on STEM Education in Early Childhood

Order	Journals	*f*	%
1	Early Childhood Education Journal	6	22.22
2	Early Education and Development	2	7.41
3	Computers in Human Behavior	1	3.70
4	Disability and Rehabilitation: Assistive Technology	1	3.70
5	Early Child Development and Care	1	3.70
6	Early Years International Research Journal	1	3.70
7	European Early Childhood Education Research Journal	1	3.70
8	International Journal of Early Childhood	1	3.70
9	International Journal of Early Years Education	1	3.70
10	International Journal of Educational Research	1	3.70

11	International Journal of Science and Mathematics Education	1	3.70
12	Journal for STEM Education Research	1	3.70
13	Journal of Applied Developmental Psychology	1	3.70
14	Journal of Children and Media	1	3.70
15	Journal of Early Childhood Research	1	3.70
16	Journal of Early Childhood Teacher Education	1	3.70
17	Journal of Experimental Child Psychology	1	3.70
18	Learning, Culture, and Social Interaction	1	3.70
19	Media Psychology	1	3.70
20	Research in Science Education	1	3.70
21	Science Activities Projects and Curriculum Ideas in STEM Classrooms	1	3.70
22	Thinking Skills and Creativity	1	3.70
23	Early Childhood Research Quarterly	1	3.70

Table 1 shows that 29 relevant articles were published in 23 different journals. Six of these articles were published in the *Early Childhood Education Journal*. A single relevant article was published in each of the other journals.

Country and Subject Distributions of the Articles

The results of the analysis of the relevant articles about STEM in early childhood in terms of country and subject are given in Table 2.

Table 2

Subject Distribution of Relevant Articles from Different Countries

Subjects	Countries where the research was conducted								
	USA	Australia	Turkey	China	Canada	Sweden	Taiwan	UK & USA	Total
Gender studies	3	3						1	7
Models design	3								3
STEM activities			2						2
Using robotics					1	1			2
Technology integration	1	1							2
Art integration						1			1
Analysing articles				1					1
Creativity process			1						1
Engineering design process			1						1
STEM identity	1								1
Integration				1					1
Teaching Mathematics				1					1
Intellectual Habits		1							1
Professional development		1							1

Program development					1				1
Self-efficacy							1		1
Teaching STEM	1								1
Thinking styles			1						1
Total	9	6	5	3	2	2	1	1	29

Table 2 outlines the relationships between the countries where these articles were published and the STEM topics considered. The United States is the country producing the highest number of relevant articles, with nine articles. Australia is in the second place with six articles, and Turkey is in the third place with five articles. Furthermore, most of the articles were written on the topic of gender in STEM education (*f*=7). This is followed by the topics of STEM models, activities, and robots.

Method and Duration of the Research

The applied methods and durations of research of the analyzed articles on STEM education in early childhood from 2016-2021 are given in Table 3.

Table 3

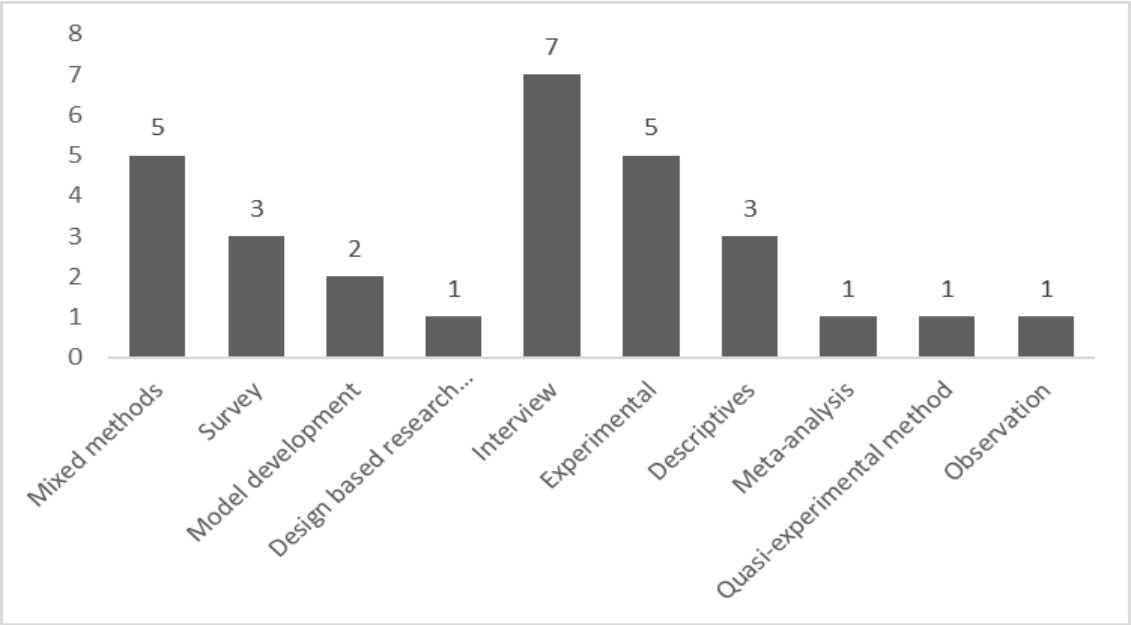
Methods and Research Durations of the Considered Articles

Methods	Qualitative	*f*	Quantitative	*f*	Mixed	*f*	Others	*f*
Time	1 session	6	8 weeks	3	52 weeks	1	104 weeks (design-based)	1
	104 weeks	1	12 weeks	1	10 weeks	1	72 weeks (meta-approach)	1
	52 weeks	1	4 weeks	1	8 weeks	1	6 weeks (Model development)	1
	2 weeks	2	1 week	1	1 session	2	1 session (workshops)	1
			1 session	4				
Total		10		11		5		4

Table 3 provides an overview of the timeframes used in early childhood STEM education research, categorized by research methods. The majority of qualitative studies (*f*=10) were conducted within shorter timeframes, such as one session or two weeks. Quantitative research (*f*=11) extended over various durations, most commonly eight weeks. Mixed methods (*f*=5) research exhibited diverse timeframes, ranging from 1 session to 104 weeks, with some involving design-based approaches. The category "Others" (*f*=4) includes various timeframes, with one study lasting for 72 weeks using a meta-approach and others involving workshops or one session.

Graph 2

Research Design of the Considered Articles



Graph 2 outlines the various research methodologies employed in early childhood STEM education studies. Mixed methods (f=5) emerge as a popular choice, reflecting the comprehensive approach researchers take to investigate this complex field. Interviews (f=7) are widely used, highlighting the significance of gathering firsthand insights from key stakeholders such as educators, parents, and children. Experimental methods (f=5) indicate a focus on empirical research to assess the impact of STEM education interventions. Surveys (f=3) and descriptives (f=3) demonstrate an interest in collecting quantitative data and statistical analyses. Model development (f=2) suggests efforts to create innovative frameworks for early childhood STEM education, and the presence of meta-analysis (f=1) shows an inclination toward synthesizing existing research findings.

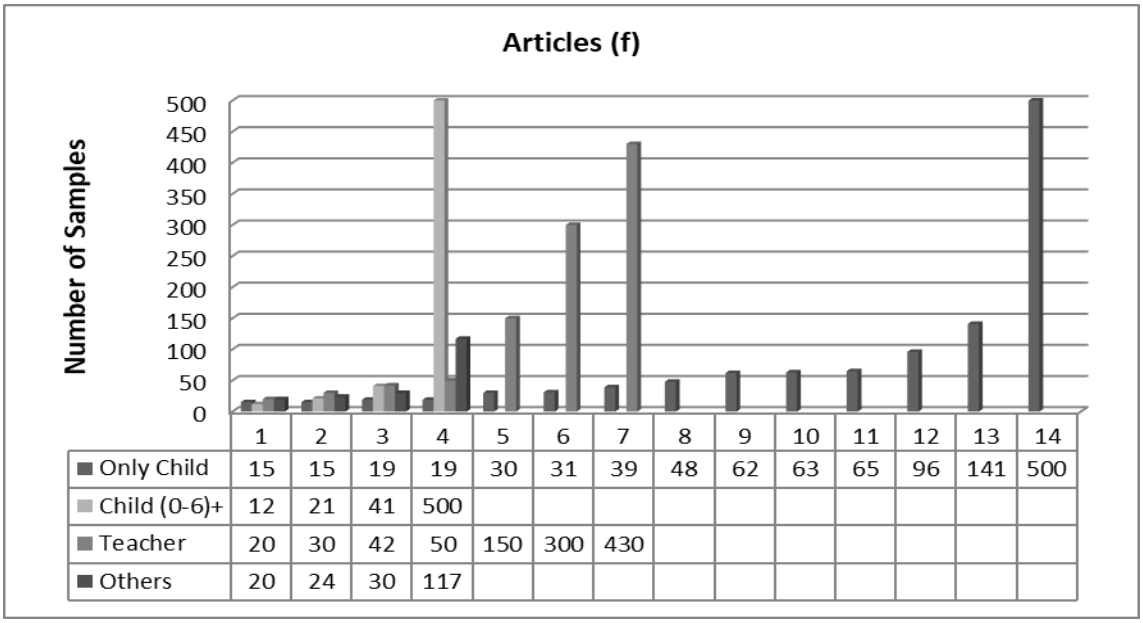
Sample Types and Sample Numbers

The sample types and numbers of the considered articles on STEM education in early childhood are illustrated in Graph 3.

Graph 3 indicates that 14 articles focused solely on children in the 0-6 age group as their sample. These studies involved a minimum of 15 and a maximum of 500 students. Additionally, four articles included both children and teachers or parents, with participant numbers ranging from 12 to 500. Seven articles centered on teachers as the sample, while four others involved different groups, such as experts, or were based on the analysis of observations of children aged 0-6 years. These articles included a minimum of 20 and a maximum of 430 participants.

Graph 3

Sample Types and Numbers of the Considered Articles



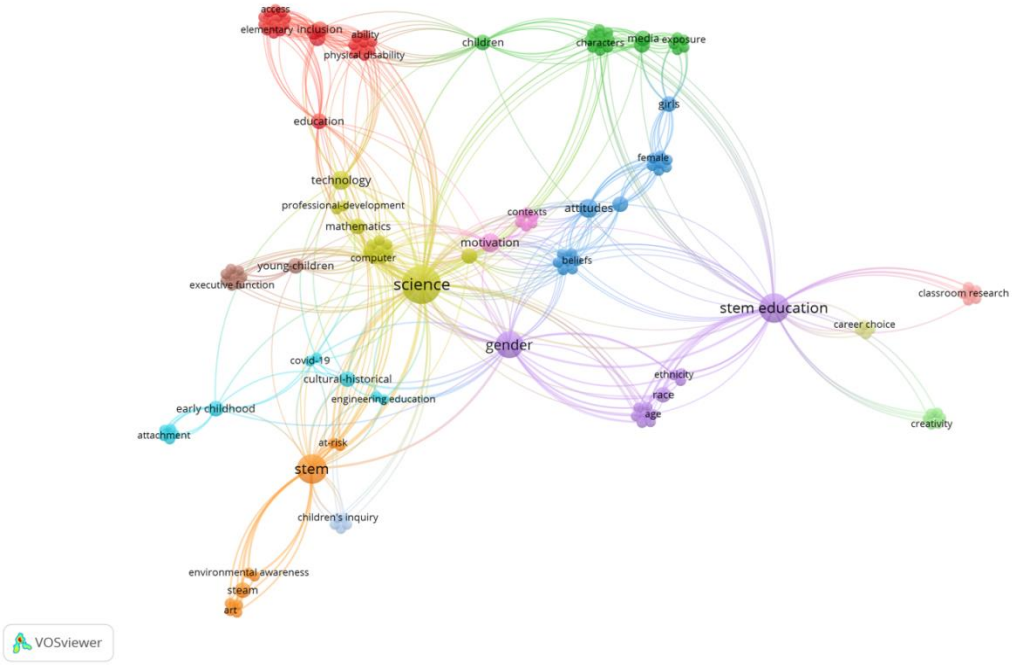
	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Only Child	15	15	19	19	30	31	39	48	62	63	65	96	141	500
Child (0-6)+	12	21	41	500										
Teacher	20	30	42	50	150	300	430							
Others	20	24	30	117										

Keywords Used in the Articles

When the keywords of the considered articles published between 2016 and 2021 were examined, the concept cloud in Figure 1 was obtained.

Figure 1

The Concept Cloud is Formed by the Keywords of the Considered Articles



When Figure 1 is examined, it is evident that key concepts such as science, STEM education, childhood, and early childhood education are particularly prevalent. These keywords are natural for the research topics at hand. However, concepts such as

gender, professional development, and science education also emerge among the keywords. In more recent research, COVID is also a topic of interest.

Data Collection Tools Used in the Articles

When the data collection tools used in the considered articles published between 2016 and 2021 were examined, the concept cloud in Figure 2 was obtained.

Figure 2

Concept Cloud of the Data Collection Tools of the Considered Articles

When Figure 2 is examined, it is evident that interviewing and observations are most prevalent among the data collection tools. In addition to these, document reviews, interview forms, and teacher interviews are also used. It can be noted that there are differences in the data collection tools used in these articles.

Relationships of the Results and Aims of the Articles

Data regarding the activities used in the considered articles published between 2016 and 2021 are presented in Table 4.

Table 4

Relationships of the Results and Aims of the Articles

Aims	Results
Model Development	1. A model was developed for bilingual students by conducting a workshop under the supervision of education leaders. 2. It was determined that the model prepared with STEM activities gave students a high level of understanding.

Effects of Gaming	1. With the support of the PlayWorld game, it was determined that girls contribute to the formation of STEM identities. 2. It was determined that technology-integrated games contribute to the development of STEM skills. 3. With technology-integrated games, it was determined that girls' interest and self-efficacy are strengthened as much as that of boys. 4. It was determined that girls participate as much as boys on playgrounds where it is difficult for girls to enter.
Program Development	1. A STEM program with parent participation was developed. 2. Expert opinions about the Little Scientists program, which was implemented for two years, were analyzed. 3. The Makerspace program was piloted for the formation of students' STEM identities. 4. A STEM program was developed for early childhood. 5. As a result of the pilot application of a STEM robotics program, it was determined that children with disabilities enjoyed the activities and gained knowledge. 6. How digital and analog programming and the use of decontextualized language may influence teachers' and children's communication about robotics and STEM in preschool was understood.
Activity Effects	1. It was determined that girls had difficulties during engineering applications and their movements were limited. 2. Children's STEM identities are formed in environments such as science museums, zoos, and aquariums. In addition, it was found that boys stated that they were better than girls in STEM subjects. 3. STEM activities increased children's creative thinking and problem-solving skills. 4. As a result of STEM activities, an increase in mathematical skills was observed, especially among middle-level and low-level students. 5. By showing children videos on STEM disciplines and gender, it was determined that the selective confidence of male characters was equal to that of boys and that of girls to male and female characters. 6. A STEM activity for egg transport was designed and explained in detail. 7. It was revealed that the effect on children's creativity was positive and permanent. 8. They showed that Makerspace environments that support 3D design can be used to encourage preschool children's STEM literacy. 9. It showed that parents' attitudes and beliefs towards science and mathematics were related to children's evaluations. Differences in reasoning and gender of the target were also identified.
Data Collection	1. Expert opinion was obtained to determine STEM concepts, ideas, and beliefs. 2. Attitudes and opinions of teachers towards STEM practices in kindergartens were obtained. 3. The STEM self-efficacy of teacher candidates was analyzed. 4. The teachers' strategies and methods in the STEM activities process were determined. 5. A strategy was determined for teachers to understand and teach STEM and to conduct art activities using digital technologies. 6. A workshop was held to increase teachers' STEM activities with Ramps & Pathways events. 7. An analysis of TV programs providing STEM education was conducted. In general, it was determined that the programs tended to exhibit more egalitarian structures. 8. An analysis of 24 articles published on early childhood was conducted.

The connections between the results and aims of the articles were examined, as shown in Table 4. In terms of the results of the methods and activities applied in these articles, the aims were classified into five groups. Nine articles aimed to understand the effects of activities, eight aimed at data collection, six explored program development, four considered the effect of playing games, and two aimed at model development. In articles investigating the effects of activities, goals included developing children's thinking, mathematical skills, and creativity in the STEM field, sometimes in the context of gender. In the articles aimed at data collection, information about early childhood was collected from teachers, experts, and sources such as TV programs. In the articles on program development, STEM programs were developed with Little Scientists, Makerspace, robotic applications, and digital or analog applications. In the articles addressing the effect of playing games, it was concluded that the STEM gender roles of girls generally matured. In the articles on model development, a model was designed for bilingual children and expected to have a higher level of understanding.

Discussion

As a result of this research and analysis of articles on STEM in early childhood, it is clear that researchers have had an interest in the STEM field ever since the concept of STEM education emerged in the late 1990s and early 2000s. With the publication of more articles on STEM education, scoping reviews continued to grow in number in the following years. For example, Thibaut et al. (2018) analyzed 23 STEM articles as a result of ERIC and Web of Science research. Martín-Páez et al. (2019) analyzed 27 articles on STEM education published between 2013 and 2018 by searching in the Web of Science database. Özkaya (2019) conducted a scoping review of STEM education studies published between 1992 and 2017. Wan et al. (2020) analyzed 24 articles using the same keywords in databases such as EBSCOhost and ERIC for studies published between 2009 and 2020. Sırakaya and Sırakaya (2020) studied research on augmented reality in STEM education from 1980 to early 2019 and analyzed 42 articles from ERIC, ProQuest, EBSCO, ScienceDirect, and Web of Science. In the study conducted by Li et al. (2020), a total of 798 articles from 45 journals on STEM subjects or disciplines published between 2000 and 2018 were examined. Takeuchi et al. (2020) analyzed a total of 143 interdisciplinary STEM articles published between 2007 and 2017. As a result of these scoping reviews performed over the years, it can be said that there has been an increase in both the number and variety of publishers and articles addressing STEM topics. Novia et al. (2021) recently examined 260 articles on educational games in the STEM field published between 2010 and 2020. Arifin and Mahmud (2021) reviewed six databases (SCOPUS, Science Direct, ERIC, Taylor & Francis, Web of Science, and Springer) and analyzed seven articles on design-oriented STEM education. In addition, in the study undertaken by Jin (2021), 24 articles on science and STEM education published between 2011 and 2020 were analyzed in a systematic review. When we look at the reviews published in 2021, it is clear that this research was performed based on articles provided by specific publishers. Considering both the journals and publishers examined within the scope of this research, it is also evident that there has been an increase in the number of articles published on the subject of STEM education from 2016 to 2021. The scoping review further indicates a rise in the number of publications in STEM education journals in recent years. The publication of relevant studies in many different journals also supports this finding. An exception is

the scoping review conducted by Denton and Borrego (2021), who examined 42 articles on STEM education at the K-12 level in their study.

When the considered articles are examined, it is evident that they were published based on research conducted in the USA, Turkey, and Australia in that order of prevalence. Thus, consistent with previous reviews, the majority of publications on STEM education research were made by authors from the USA, where STEM and STEAM education first arose. Martín-Páez et al. (2019), Jin (2021), Le Thi Thu et al. (2021), and Han et al. (2021) all reported the most publications being made by authors from the USA in their analysis of articles on STEM education. At the same time, authors from Australia, Canada, Taiwan, and some other parts of Asia have become more active in this field over the past few years (Li et al., 2020). Although these previous findings are consistent with the present study's findings, there has been a trend toward increasing numbers of STEM education publications in Turkey in recent years. Regarding topics, our research revealed that the largest number of articles addressed gender in STEM education. Other subjects of particular interest were STEM models, activities, and robots. Similarly, Tselegkaridis and Sapounidis (2021) stated that robotic applications were common in STEM education studies. Martín-Páez et al. (2019) determined that articles largely focused on STEM education, STEM literacy, and STEM curricula. Thus, while articles on STEM education and STEM curricula have historically been dominant in literature reviews, gender in STEM education has grown more prominent as a research topic in recent years, according to our analysis.

In the considered articles on STEM education in early childhood, qualitative methods were mostly preferred, and most studies were performed with one-time applications. In addition to qualitative methods, experimental, mixed, and screening methods were also used. While some considered articles presented the results of one-time applications, others had research durations extending to 104 weeks. Similarly, in the analysis of articles on STEM education, Jin (2021) reported a majority of qualitative methods with application periods ranging from 2 weeks to 1 year. Similarly, in the analysis of STEM articles by Jayarajah et al. (2014), it was concluded that qualitative studies were in the majority. These findings support the conclusions of the present work. In contrast, Martín-Páez et al. (2019) reported that a majority of STEM education articles employed mixed methods. Seven and Uçar (2020) identified quantitative methods as being predominant in an examination of theses on preschool education. It is clear that qualitative methods have been generally preferred in recent years for studies of STEM education in preschool, in addition to other studies conducted with methods such as quantitative and mixed methods. Both short-term and long-term studies were carried out within this framework.

In terms of sample types and numbers, the majority of articles focused on sample groups consisting solely of children aged 0-6 years. These articles were prepared with a minimum of 15 students and a maximum of 500 students. Other publications have considered teachers, parents, experts, and previously published articles addressing students aged 0-6 years. In this group of articles, sample sizes range from 12 to 430. Since the present study specifically considered STEM articles addressing the age group of 0-6 years, the samples predominantly comprised preschool students and stakeholders (teachers, parents, experts). For this reason, no similar study could be identified in the literature. However, there are related studies supporting these findings. For example,

Martín-Páez et al. (2019) stated that very few publications address STEM education for preschool children. Ahi and Kıldan (2013), Çifçi and Ersoy (2019), and Şahin and Bartan (2017) determined that sample groups generally comprised children in their analyses of theses written about the preschool period. In addition, when Kiremit (2019) and Karoğlu and Esen Çoban (2019) analyzed publications on language development in children, they determined that the majority of the sample groups comprised children. This is thought to be related to the research subjects at hand. Sırakaya and Sırakaya (2020) and Le Thi Thu et al. (2021) determined that their considered articles were mostly written with secondary school students as the sample populations, while Jayarajah et al. (2014) determined that the majority of participants were in the age group of 12-24 in their analysis of STEM education articles published between 1999 and 2013. As can be seen here, STEM education studies have mostly been conducted for middle school students or the age group of 12-24 years (i.e., high school and university) rather than preschool. When studies are examined in terms of the sizes of samples, Jin (2021) determined the number of participants to range between 24 and 42 in the study. Çifçi and Ersoy (2019) and Sırakaya and Sırakaya (2020) found that the majority of considered articles were based on research conducted with 31-100 participants. These findings are similar to those of the present study.

In the articles published between 2016 and 2021, the keywords "STEM," "STEM education," "childhood," and "early childhood education" particularly stand out. These keywords are logical considering the topics of the articles. Similarly, Oğurlu and Çayır (2014) analyzed the key concepts explored in publications on gifted people. They determined that the key concepts most often considered were related to the subject at hand, such as intelligence or science and art centers. In the present study, in addition to the key concepts directly related to the subject, concepts such as gender, professional development, and science education are also observed. Furthermore, in the most recent studies, the COVID-19 pandemic is mentioned. Notably, interviews and observations are primarily used as data collection tools in articles on STEM education in early childhood. In addition, tools such as document reviews, interview forms, and teacher interviews may be mentioned. Thus, there are differences in the data collection tools used in the considered articles, varying depending on sample types and study methods. Sırakaya and Sırakaya (2020) found that the majority of their considered studies involved secondary school students, and the data were accordingly collected mainly through achievement tests, screenings, and interviews. Although interviews and screening techniques were similarly observed in the present research, there were no findings for achievement tests.

The activities or practices applied in the considered articles were classified into five groups according to their purposes. Generally, the majority of the aims included determining the STEM identities of girls with games; developing child-specific or parent-participation programs such as robotics, Little Scientists, or Makerspace; applying STEM activities to determine their effects on assorted variables; and collecting data, primarily from teachers, on the application of STEM education. Similarly, Wan et al. (2020) reported articles addressing the effects of digital games and robotic applications on assorted variables and STEM education with the opinions of teachers, parents, and experts. These research findings are similar to the results of the present study in terms of classifying the aims of the reviewed articles. Takeuchi et al. (2020)

found that most of the considered articles focused on students' STEM career plans. Thibaut et al. (2018) stated that the articles they considered were predominantly about STEM integration. These findings are compatible with those of the present study as well. The effects of STEM activities on variables such as selective confidence, mathematical skills, and problem-solving have been examined in the relevant literature. In their study, Ha et al. (2020) evaluated scientific results on STEM education, including engineering and computing education. Jin (2021) reported that among the aims of the articles he analyzed, improving scientific inquiry skills, developing interest in and positive attitudes towards STEM subjects, and building STEM identities were the top ones. Although these findings do not exactly overlap with those of the present work, there are similarities. Sırakaya and Sırakaya (2020) classified the purposes and learning outcomes of the articles they considered and reported a focus on the effect of learning, the outcomes of training, and interactions. In addition, Martín-Páez et al. (2019) concluded that their considered articles were mostly about cognitive development and attitudes. In the analysis of STEM articles published between 2010 and 2015, McDonald (2016) concluded that the role of teachers in quality learning is important as they can increase students' interest in and positive attitudes towards STEM subjects with effective pedagogical formation. Thus, these studies mostly focus on analyzing changes in interests, attitudes, and cognitive development based on STEM education in different sample groups.

In general, it can be asserted that studies on STEM education in early childhood show great promise. Involving children in STEM activities with robotics applications, games, activities, models, or program development contributes to the field. Moreover, better STEM activities can be planned with studies conducted based on the opinions of teachers and experts. The engagement of parents in this process is considered particularly promising for future studies and the development of successful STEM education. These findings reflect that achievements in STEM activities will have impacts at home, at school, on the streets, and in all areas of daily life. Based on the assumptions and limitations of studies conducted to date, new studies can be undertaken, and the impact of STEM education can be increased. In addition, it can be noted that studies aimed at providing a gender balance in STEM education have become widespread. There may be classroom interactions in which girls and boys will participate at similar rates. Finally, the quality of education should be increased for individual students by taking economic and intellectual variables into account. The effects of using STEM disciplines for the acquisition of daily life skills have become particularly evident.

Conclusion and Recommendations

Although there has been an increase in the number of articles on preschool STEM education published between 2016 and 2021, this is still a relatively new and growing field compared to STEM education studies for different age groups. For this reason, it is recommended that STEM education researchers focus more on STEM education in the preschool period.

Upon reviewing the publication of articles on STEM education in early childhood, it has been determined that there is a stronger tendency to conduct STEM education research in developed countries than in developing countries. Developing

countries should be supported in this regard, and the STEM subjects studied for the preschool age group should be diversified.

Qualitative methods were primarily used in the articles analyzed in this research, although other methods were also observed. It is recommended to use the operational research method for both close monitoring and individual development at the preschool level. In addition, although both short-term and long-term studies were observed in terms of the duration of research, more programmatic or achievement-oriented studies should be conducted.

In this examination of articles addressing the age group of 0-6 years, it was noted that the sample groups also included relevant stakeholders. The number of studies addressing these stakeholders should be increased. In addition, studies with differing sample sizes were observed, and it should not be forgotten that it is important that the number of samples be appropriate for the selected method.

While it was generally clear that articles on STEM education in early childhood use keywords related to their subjects, the concepts of gender, professional development, and science education were particularly prevalent. The use of different concepts enriches the field. It is suggested that a common standard be embraced due to differences in keywords, such as "STEM" in some studies and "STEM education" in others.

The prevalent use of interviews and observations as data collection tools was an expected finding due to the fact that these articles were largely based on qualitative methods and the majority of participants, i.e., children aged 0-6, were illiterate. It is recommended to use multidimensional data collection tools simultaneously to evaluate studies conducted with children.

When the STEM activities for preschoolers in the considered articles were examined, it was determined that different activities are carried out to increase the effects of STEM education. These activities designed for the age group of 0-6 generally focus on games, teachers' roles, or program and model development. For such activities, it is necessary to determine factors such as the role of sample groups in the process and their impact on the process. In other words, there is a need for studies focusing on the evaluation dimension of STEM education in preschool.

Limitations

The study focused on articles published by specific publishers (Springer, Taylor & Francis, Elsevier, and SAGE). This approach may not capture the entire body of relevant literature in the field of preschool STEM education, potentially leading to publication bias. The study analyzed articles published from January 2016 to June 2021, which may exclude earlier research that could provide valuable historical context and insights into the development of STEM education in the preschool period. The criterion of analyzing articles addressing the age group of 0-6 years led to the exclusion of 25 articles. This selection criterion may introduce subjectivity and potentially exclude relevant studies with slightly different age group classifications. Scoping reviews, by design, provide a broad overview of the literature but may lack the depth of analysis and critical appraisal seen in systematic reviews. This limitation should be acknowledged, as it may impact the level of detail in the review.

Compliance with Ethical Standards

This research has not been done with humans, animals, or any other living things. The research does not contain elements that would threaten the health of any living thing. In the research, previous academic publications on a specific subject were analyzed. Information and data belonging to others were not used in the research, and references were made to the publications used. There was no potential conflict of interest in the study, neither between the authors nor with anyone else. This research is not a part of a study; it shows integrity in itself. The research was not sent to any journal other than this journal in accordance with the publication ethical rules.

Acknowledgements

No funding was received from any institution for the preparation of the article. In addition, the contribution of the authors in the preparation of the article is equal.

Statement of Responsibility

All parts of the article were co-written by the authors. All sections such as Introduction, Method, Results, comment, discussion, and conclusion were written together. All expenses related to the article were made by the authors.

Conflicts of Interest

There is no conflict of interest since the authors or the article do not have a financial, commercial, legal or professional relationship with any person or organization.

Author Bios:

İsmail Dönmez received his undergraduate education at Çanakkale Onsekiz Mart University, his graduate education at Eskişehir Osmangazi University, and his doctoral education at Gazi University. Dönmez works on topics such as learning environments, STEM education, giftedness and self-examination. He is an ambassador for European School-Net's Scientix Project and a Texas Instruments T3 European instructor. He is a member of Hacettepe University STEM & Maker Lab team. He took part in many TUBITAK projects and European Commission FP7 projects (eg SAILS, MaScil). He is also the editor of the Journal of STEAM Education (J-STEAM). He is the chairman of the board of directors of the STEAM Education Research Association. He is still working as a faculty member in the Department of Child Care and Youth Services at Muş Alparslan University.

Salih Gülen was born in 1984 in Iğdır, and he completed his primary, secondary and high school education in Iğdır and his university education at Van Yüzüncü Yıl University. He did his master's degree on technology integration in education, his doctorate on STEM education and integration of STEM education into secondary school curriculum. Gülen, who received his doctorate and associate professorship in science education, has made many academic publications in national and international indexes on technology use and STEM education. He has worked as an administrator and teacher in the provinces of Van, Tekirdağ, Samsun and Muğla for many years, and has been working as a faculty member at Muş Alparslan University since 2018.

References

Ahi, B., & Kıldan, A. O. (2013). An overview of postgraduate thesis within the field of preschool education in Turkey (2002-2011). *Journal of Mehmet Akif Ersoy University Faculty of Education, 13*(27), 23-46.

Aladé, F., Lauricella, A., Beaudoin-Ryan, L., & Wartella, E. (2016). Measuring with murray: Touchscreen technology and preschoolers' STEM learning. *Computers in Human Behavior, 62*, 433-441. https://doi.org/10.1016/j.chb.2016.03.080

Aldemir, J., & Kermani, H. (2016). Integrated STEM curriculum: Improving educational outcomes for head start children. *Early Child Development and Care, 187*(11), 1694-1706. https://doi.org/10.1080/03004430.2016.1185102

Arifin, N. R., & Mahmud, S. N. D. (2021). A systematic literature review of design thinking application in STEM integration. *Creative Education, 12*(7), 1558-1571, https://doi.org/10.4236/ce.2021.12711810.4236/ce.2021.127118

Arksey, H., & O'Malley, L. (2005). Scoping studies: towards a methodological framework. *International Journal of Social Research Methodology, 8*(1), 19-32.

Armstrong, R., Hall, B. J., Doyle, J., & Waters, E. (2011). Scoping the scope'of a cochrane review. *Journal of public health, 33*(1), 147-150.

Ata-Aktürk, A., & Demircan, H. Ö. (2021). Supporting preschool children's STEM learning with parent-involved early engineering education. *Early Childhood Education Journal, 49*(4), 607-621. https://doi.org/10.1007/s10643-020-01100-1

Brenneman, K., Lange, A., & Nayfeld, I. (2019). Integrating STEM into preschool education; designing a professional development model in diverse settings. *Early Childhood Education Journal, 47*(1), 15-28. https://doi.org/10.1007/s10643-018-0912-z

Brophy, S., Klein, S., Portsmore, M., & Rogers, C. (2008). Advancing engineering education in P-12 classrooms. *Journal of Engineering Education, 97*(3), 369–387. https://doi.org/10.1002/j.2168-9830.2008. tb00985.x

Bybee, R. W. (2013). *The case for STEM education; challenges and opportunities*. National Science Teachers Association Press.

Campbell, C., Hobbs, L., Millar, V., RagabMasri, A., Speldewinde, C., Tytler, R., & van Driel, J. (2020). *Girls' future – Our future. The Inver Gowrie Foundation STEM report 2020 update*. Invergowrie Foundation

Çetin, M. & Demircan, H. (2020). STEM education in early childhood. *İnönü Üniversitesi Eğitim Fakültesi Dergisi, 21*(1), 102-117. https://doi.org/10.17679/inuefd.437445

Çifçi, M., & Ersoy, M. (2019). Trends of research in the field of preschool education: a content analysis. *Cumhuriyet International Journal of Education, 8*(3), 862-886. http://dx.doi.org/10.30703/cije.581302

Colquhoun, H., Levac, D., O'Brien, K., Straus, S., Tricco, A., Perrier, L., Kastner, M., & Moher, D. (2014). Scoping reviews: Time for clarity in definition, methods, and reporting. *Journal of Clinical Epidemiology, 67*(12), 1291-1294. https://doi.org/10.1016/j.jclinepi.2014.03.013

DeJarnette, N. K. (2018). Implementing STEAM in the early childhood classroom. *European Journal of STEM Education, 3*(3), 18. https://doi.org/10.20897/ejsteme/3878

Denton, M., & Borrego, M., (2021). Funds of knowledge in STEM education: A scoping review. *Studies in Engineering Education, 1*(2), 71-92, https://doi.org/10.21061/see.19

Early Childhood STEM Working Group. (2017). *Early STEM matters providing high-quality STEM experiences for all young learners: A policy report by the early childhood STEM working group*. Chicago, IL: University of Chicago and the Erikson Institute.

Greenfield, D., Jirout, J., Dominguez., X., Greenberg, A., Maier, M., & Fuccillo, J. (2009). Science in the preschool classroom: A programmatic research agenda to improve science readiness. *Early Education & Development, 20*, 238–264. https://doi.org/10.1080/10409280802595441

Guzey, S. & Harwell, M. (2016). Building up STEM: An analysis of teacher-developed engineering design-based STEM integration curricular materials. *Journal of Pre-College Engineering Education Research (J-Peer), 6*(1). https://doi.org/10.7771/2157-9288.1129

Ha, C. T., Thao, T. T. P., Trung, N. T., Huong, L. T. T., Dinh, N. V., & Trung, T., (2020). A bibliometric review of research on STEM education in ASEAN: science mapping the literature in scopus database, 2000 to 2019. *EURASIA Journal of Mathematics, Science and Technology Education, 16*(10), 1-11, https://doi.org/10.29333/ejmste/8500

Hadzigeorgiou, Y. (2002). A study of the development of the concept of mechanical stability in preschool children. *Research in Science Education, 32*(3), 373–391. https://doi.org/10.1023/A:1020801426075

Hammack, R., & Ivey, T. (2017). Examining elementary teachers' engineering self-efficacy and engineering teacher efficacy. *School Science and Mathematics, 117*(1-2), 52–62. https://doi.org/10.1111/ssm.1220

Han, J., Kelley, T. R., Mentzer, N., & Knowles, J. G. (2021). Community of practice in integrated STEM education: A systematic literature review. *Journal of STEM Teacher Education, 56*(2), 62-80.

Hemingway, P., & Brereton, N. (2009). *What is a systematic review*? http://www.whatisseries.co.uk/whatis/

Jang, H. (2016). Identifying 21st century STEM competencies using work place data, *Journal of Science Education and Technology, 25*(2), 284-301. https://doi.org/10.1007/s10956-015-9593-1

Jayarajah, K., Saat, R. M., & Rauf, R. A. A. (2014). A review of science, technology, engineering & mathematics (STEM) education research from 1999–2013: A Malaysian perspective. *Eurasia Journal of Mathematics, Science & Technology Education, 10*(3), 155-163. https://doi.org/10.12973/eurasia.2014.1072a

Jin, Q. (2021). Supporting indigenous students in science and STEM education: A systematic review. *Education Sciences, 11*(9), 555, https://doi.org/10.3390/educsci11090555

Karampelas, K. (2021). Trends on science education research topics in education journals. *European Journal of Science and Mathematics Education, 9*(1), 1-12. https://doi.org/10.30935/scimath/9556

Karaşah Çakıcı, Ş., Kol, Ö., & Yaman, S. (2021). The effects of STEM education on students' academic achievement in science courses: A meta-analysis. *Journal of Theoretical Educational Science, 14*(2), 264-290. https://doi.org/10.30831/akukeg.810989

Karoğlu, H., & Esen Çoban, A. (2019). Investigation of graduate theses on the language development in preschool education in Turkey. *Journal of Erzincan University Faculty of Education, 21*(1), 211-229. https://doi.org/10.17556/erziefd.432657

Katz, L. (2010, May). *STEM in the early years*. Paper presented at the SEED 2010: STEM in early education and development conference. Retrieved from http://ecrp.uiuc.edu/beyond/seed/

Kayan-Fadlelmula, F., Sellami, A., Abdelkader, N., & Umer, S. (2022). A systematic review of STEM education research in the GCC countries: Trends, gaps and barriers. *International Journal of STEM Education, 9*(1), 1-24. https://doi.org/10.1186/s40594-021-00319-7

Kiremit, R. F. (2019). An overview of postgraduate theses within the field of language development at early childhood education in Turkey. *Eskişehir Osmangazi University Journal of Social Sciences, 20*, 159-174. https://doi.org/10.17494/ogusbd.548314

Kürklü, S., (2019). Bibliometric Analysis of Research on Bioethics; WOS Example. *Turkish Journal of Bioethics, 6*(3), 87-99.

Lange, A. A., Brenneman, K., & Mano, H. (2019). *Teaching STEM in the preschool classroom: Exploring big ideas with 3-to 5-year olds*. Teachers College Press.

Le Thi Thu, H., Tran, T., Trinh Thi Phuong, T., Le Thi Tuyet, T., Le Huy, H., Vu Thi, T. (2021). Two decades of STEM education research in middle school: A bibliometrics analysis in scopus database (2000–2020). *Education Science, 11*, 353. https://doi.org/10.3390/educsci11070353

Leung, W. M., & Xinyun, H. U. (2019). We want STEM: Exploring digital toys in a Hong Kong kindergarten. *Journal of Education and Human Development, 8*(4), 82-93. https://doi.org/10.15640/jehd.v8n4a11

Li, Y., Wang, K., Xiao, Y., &Froyd, J. E. (2020). Research and trends in STEM education: A systematic review of journal publications. *International Journal of STEM Education, 7*(11), 2-16. https://doi.org/10.1186/s40594-020-00207-6

Manowaluilou, N., & Nilsook, P. (2023). Career awareness link age strategies to support learning career education and STEM education. *Kasetsart Journal of Social Sciences, 44*(1), 199-208. https://doi.org/10.34044/j.kjss.2023.44.1.22

Martín-Páez, T., Aguilera, D., Perales-Palacios, F. J., & Vílchez-González, J. M., (2019). What are we talking about when we talk about STEM education? A review of literature. *Science Education, 103*, 799–822. https://doi.org/10.1002/sce.21522

Mcdonald, C. V. (2016). STEM Education: A review of the contribution of the disciplines of science, technology, engineering and mathematics. *Science Education International, 27*(4), 530-569.

Morgan, P. L., Farkas, G., Hillemeier, M. M., & Maczuga, S. (2016). Science achievement gaps begin very early, persist, and are largely explained by modifiable factors. *Educational Researcher, 45*(1), 18-35. https://doi.org/10.3102/0013189X16633182

Morrison, J. (2006). *TIES STEM education monograph series, Attributes of STEM education*. MD: TIES

Moule, P., Ward, R., & Lockyer, L. (2010). Nursing and healthcare students' experiences and use of e-learning in higher education. *Journal of Advanced Nursing, 66*(12), 2785-2795. https://doi.org/10.1111/j.1365-2648.2010.05453.x

Munn, Z., Peters, M. D., Stern, C., Tufanaru, C., McArthur, A., & Aromataris, E. (2018). Systematic review or scoping review? Guidance for authors when choosing between a systematic or scoping review approach. *BMC medical research methodology, 18*(1), 1-7.

Murphy, S., MacDonald, A., Danaia, L., & Wang, C. (2018). An analysis of Australian STEM education strategies. *Policy Futures in Education, 17*(2), 122-139. https://doi.org/10.1177/1478210318774190

National Academy of Engineering and National Research Council [NSTA]. (2014). *STEM integration in K 12 education: Status, prospects, and an agenda for research*. National Academies Press.

National Research Council [NRC]. (2010). *Exploring the intersection of science education and 21st century skills: A workshop summary.* National Academies Press.

National Research Council [NRC]. (2011). *Successful K-12 STEM education: Identifying effective approaches in science, technology, engineering, and mathematics*. National Academies Press.

Novia, N., Permanasari, A., & Riandi, R., (2021). Research on educational games in STEM area 2010-2020: A bibliometric analysis of literature. *International Conference on Mathematics and Science Education*, https://doi.org/10.1088/1742-6596/1806/1/012209

Oğurlu, Ü., & Çayır, Ş. (2014). Analysis of keywords used in research about gifted. *Journal of Gifted Education Research, 2*(2), 72-85.

Özkaya, A. (2018). Bibliometric analysis of the studies in the field of mathematics education. *Educational Research and Reviews, 13*(22), 723-734. http://dx.doi.org/10.5897/ERR2018.3603

Özkaya, A. (2019). Bibliometric analysis of the publications made in STEM education area. *Bartın University Journal of Faculty of Education, 8*(2), 590-628 https://doi.org/10.14686/buefad.450825

Peters, M. D. J., Godfrey, C. M., Khalil, H., McInerney, P., Parker, D., & Soares, C. B., (2015). Guidance for conducting systematic scoping reviews. *International Journal of Evidence-Based Healthcare 13*(3), 141-146. https://doi.org/10.1097/XEB.0000000000000050

Pham, M., Rajić, A., Greig, J., Sargeant, J., Papadopoulos, A., & McEwen, S. (2014). A scoping review of scoping reviews: Advancing the approach and enhancing the consistency. *Research Synthesis Methods, 5*(4), 371-385. https://doi.org/10.1002/jrsm.1123

Phan-Le, N., Brennan, L., & Parker, L. (2022). The search for scientific meaning in mindfulness research: Insights from a scoping review. *Plos One, 17*(5), e0264924. https://doi.org/10.1371/journal.pone.0264924

Psycharis, S. (2018). STEAM in education: A literature review on the role of computational thinking, engineering epistemology and computational science. Computational STEAM pedagogy (CSP). *Scientific Culture, 4*(2), 51-72. https://doi.org/10.5281/zenodo.1214565

Şahin, G., & Bartan, M. (2017). Investigation of graduate thesis in the preschool educatıon. *The Journal of Academic Social Science Studies,* 69-84. http://dx.doi.org/10.9761/JASSS7256

Seven, S., & Uçar, S., (2020). A review of the master's theses and doctoral dissertations on inclusion in Turkey. *Ulakbilge - Journal of Social Sciences, 8*(54), https://doi.org/10.7816/ulakbilge-08-54-12

Sırakaya, M., & Sırakaya, D. A. (2020). Augmented reality in STEM education: A systematic review. *Interactive Learning Environments,* 1-14. https://doi.org/10.1080/10494820.2020.1722713

Sydon, T. & Phuntsho, S. (2022). Highlighting the importance of STEM education in early childhood through play-based learning: A literature review. *Rabsel, 22*(1). https://doi.org/10.17102/rabsel.22.1.3

Takeuchi, M. A., Sengupta, P., Shanahan, M. C., Adams, J. D., & Hachem, M. (2020). Trans disciplinarily in STEM education: A critical review. *Studies in Science Education, 56*(2), 213-253. https://doi.org/10.1080/03057267.2020.1755802

Tao, Y. (2019). Kindergarten teachers' attitudes toward and confidence for integrated STEM education. *Journal for STEM Education Research, 2*(2), 154-171. https://doi.org/10.1007/s41979-019-00017-8

Thibaut, L., Ceuppens, S., De Loof, H., De Meester, J., Goovaerts, L., Struyf, A., Boeve-de Pauw, J., Dehaene, W., Deprez, J., De Cock, M., Hellinckx, L., Knipprath, H., Langie, G., Struyven, K., Van de Velde, D., Van Petegem, P. & Depaepe, F. (2018). Integrated STEM education: A systematic review of instructional practices in secondary education. *European Journal of STEM Education, 3*(1), 02. https://doi.org/10.20897/ejsteme/85525

Tippett, C. D., & Milford, T. M. (2017). Findings from a pre-kindergarten classroom: Making the case for STEM in early childhood education. *International Journal of Science and Mathematics Education, 15*(1), 67-86. https://doi.org/10.1007/s10763-017-9812-8

Tricco, A.C., Antony, J., Zarin, W. Strifler, L., Ghassemi, M., Ivory, J., Perrier, L., Hutton, B., Moher, D., & Straus, S. E. (2015). A scoping review of rapid review methods. *BMC Medicine 13,* 224. https://doi.org/10.1186/s12916-015-0465-6

Tselegkaridis, S., & Sapounidis, T. (2021). Simulators in educational robotics: A review. *Education Sciences, 11*(11). https://doi.org/10.3390/educsci11010011

Tsujimoto, S. (2008). The prefrontal cortex: Functional neural development during early childhood. *The Neuroscientist, 14*(4), 345-358. https://doi.org/10.1177/1073858408316002

Wan, Z. H., Jiang, Y. & Zhan, Y. (2020). STEM education in early childhood: A review of empirical studies. *Early Education and Development*, https://doi.org/10.1080/10409289.2020.1814986

Wang, H. H., Moore, T. J., Roehrig, G. H., & Park, M. S. (2011). STEM integration: The impact of professional development on teacher perception and practice. *Journal of Pre-College Engineering Education Research 2*(1), 28-34. https://doi.org/10.5703/1288284314636

Yan, B., Yang, F., & Cai, N. (2021). On time effect of preschool education: Social analysis based on cucds. *Complexity*, 1-10. https://doi.org/10.1155/2021/2855542

Yang, W., Du, Y., Wu, R., & Xiang, S. (2023). Development and validation of the children's STEM habits of mind questionnaire. *Early Childhood Education Journal*. https://doi.org/10.1007/s10643-023-01451-5

Yang, W., Wu, R., & Li, J. (2021). Development and validation of the STEM teaching self-efficacy scale (STSS) for early childhood teachers. *Current Psychology, 42*(9), 7275-7283. https://doi.org/10.1007/s12144-021-02074-y

Reviewed Articles

1 Aladé, F., Lauricella, A. R., Beaudoin-Ryan, L., & Wartella, E. (2016). Measuring with Murray: Touchscreen technology and preschoolers' STEM learning. *Computers in Human Behavior*. https://doi.org/10.1016/j.chb.2016.03.080

2 Aladé, F., Lauricella, A., Kumar, Y., & Wartella, E., (2020). Who's modeling STEM for kids? A character analysis of children's STEM-focused television in the U.S. *Journal of Children and Media*. https://doi.org/10.1080/17482798.2020.1810087

3 Aldemir, J., & Kermani, H., (2017). Integrated STEM curriculum: Improving educational outcomes for head start children, *Early Child Development and Care, 187*(11), 1694-1706, https://doi.org/10.1080/03004430.2016.1185102

4 Ata-Aktürk, A., & Demircan, H.Ö. (2020). Supporting preschool children's STEM learning with parent-involved early engineering education. *Early Childhood Education Journal, 49*, 607-621. https://doi.org/10.1007/s10643-020-01100-1

5 Brenneman, K., Lange, A., & Nayfeld, I. (2019). Integrating STEM into preschool education; Designing a professional development model in diverse settings. *Early Childhood Education Journal, 47*, 15-28. https://doi.org/10.1007/s10643-018-0912-z

6 Chen, Y. L., Huang, L. F., & Wu, P. C. (2021). Preservice preschool teachers' self-efficacy in and need for STEM education professional development: STEM pedagogical belief as a mediator. *Early Childhood Education Journal, 49*, 137-147. https://doi.org/10.1007/s10643-020-01055-3

7 Counsell, S. L., & Geiken, R., (2019) Improving STEM teaching practices with R&P: increasing the full range of young children's STEM outcomes. *Journal of Early Childhood Teacher Education, 40*(4), 352-381. https://doi.org/10.1080/10901027.2019.1603173

8 Fleer, M., (2021). Re-imagining play spaces in early childhood education: Supporting girls' motive orientation to STEM in times of COVID-19. *Journal of Early Childhood Research*. https://doi.org/10.1177/1476718X20969848

9 Fleer, M., (2021). When preschool girls engineer: Future imaginings of being and becoming an engineer. *Learning, Culture and Social Interaction*. https://doi.org/10.1016/j.lcsi.2019.100372

10 Fridberg, M., & Redfors, A., (2021). Teachers' and children's use of words during early childhood STEM teaching supported by robotics, *International Journal of Early Years Education*. https://doi.org/10.1080/09669760.2021.1892599

11 Hachey, A. C., An, S. A. & Golding, D. E. (2021). Nurturing kindergarteners' early STEM academic identity through makerspace pedagogy. *Early Childhood Education Journal, 50*, 469-479. https://doi.org/10.1007/s10643-021-01154-9

12 He, X., Li, T., Turel, O., Kuang, Y., Zhao, H., & He, Q. (2021). The impact of STEM education on mathematical development in children aged 5-6 years. *International Journal of Educational Research*. https://doi.org/10.1016/j.ijer.2021.101795

13 Lindsay, S., & Hounsell, K. G. (2017). Adapting a robotics program to enhance participation and interest in STEM among children with disabilities: a pilot study, *Disability and Rehabilitation: Assistive Technology, 12*(7), 694-704. https://doi.org/10.1080/17483107.2016.1229047

14 MacDonald, A., Huser, C., Sikder, S. (2020). Effective early childhood STEM education: Findings from the little scientists evaluation. *Early Childhood Education Journal, 48*, 353-363. https://doi.org/10.1007/s10643-019-01004-9

15 Magnusson, L. O., & Bäckman, K., (2021). What is the capacity of A in the contexts of STEM?, Early Years. https://doi.org/10.1080/09575146.2021.1914557

16 Master, A., Cheryan, S., Moscatelli, A., & Meltzoff, A. N., (2017). Programming experience promotes higher STEM motivation among first-grade girls. *Journal of Experimental Child Psychology*. https://doi.org/10.1016/j.jecp.2017.03.013

17 McGuire, L., Mulvey, K. L., Goff, E., Irvin, M. J., Winterbottom, M., Fields, G. E., Hartstone-Rose, A., & Rutland, A. (2020). STEM gender stereotypes from early childhood through adolescence at informal science centers. *Journal of Applied Developmental Psychology*. https://doi.org/10.1016/j.appdev.2020.101109

18 McGuire, L., Mulvey, K. L., Goff, E., Irvin, M. J., Winterbottom, M., Fields, G. E., ... & Rutland, A. (2020). STEM gender stereotypes from early childhood through adolescence at informal science centers. *Journal of Applied Developmental Psychology, 67*, 101109. https://doi.org/10.1016/j.appdev.2020.101109

19 Mulvey, K. L., & Irvin, M. J. (2018). Judgments and reasoning about exclusion from counter-stereotypic STEM career choices in early childhood. *Early Childhood Research Quarterly, 44*, 220-230. https://doi.org/10.1016/j.ecresq.2018.03.016

20 Savinskaya, O. B. (2017). Gender equality in preschool STEM programs as a factor determining Russia's successful technological development. *Russian Education & Society, 59*(3-4), 206-216.

21 Schlesinger, M. A., & Richert, R. A., (2019) The role of gender in young children's selective trust of familiar STEM characters, *Media Psychology, 22*(1), 109-132. https://doi.org/10.1080/15213269.2017.1328311

22 Simoncini, K., & Lasen, M. (2018). Ideas about STEM among australian early childhood professionals: How important is STEM in early childhood education. *International Journal of Early Childhood, 50*, 353–369. https://doi.org/10.1007/s13158-018-0229-5

20 Stephenson, T., Fleer, M. & Fragkiadaki, G. (2021). Increasing girls' STEM engagement in early childhood: Conditions created by the conceptual playworld model. *Research in Science Education, 52*, 1243–1260. https://doi.org/10.1007/s11165-021-10003-z

23 Tao, Y. (2019). Kindergarten teachers' attitudes toward and confidence for integrated STEM education. *Journal for STEM Education Research, 2*, 154–171. https://doi.org/10.1007/s41979-019-00017-8

24 Tippett, C. D., & Milford, T. M. (2017). Findings from a Pre-kindergarten classroom: Making the case for stem in early childhood education. *International Journal of Science and Mathematics Education*, *15*(Suppl1), 67–86. https://doi.org/10.1007/s10763-017-9812-8

25 Ültay, N., & Aktaş, B., (2020). An example implementation of STEM in preschool education: Carrying eggs without breaking, *Science Activities, 57*(1), 16-24. https://doi.org/10.1080/00368121.2020.1782312

26 Üret, A., & Ceylan, R., (2021). Exploring the effectiveness of STEM education on the creativity of 5-year-old kindergarten children, *European Early Childhood Education Research Journal*. https://doi.org/10.1080/1350293X.2021.1913204

27 Wan, Z. H., Jiang, Y., & Zhan, Y., (2020). STEM education in early childhood: A review of empirical studies, *Early Education and Development*. https://doi.org/10.1080/10409289.2020.1814986

28 Yalçın, V., & Erden, Ş., (2021). The effect of STEM activities prepared according to the design thinking model on preschool children's creativity and problem-solving skills. *Thinking Skills and Creativity*. https://doi.org/10.1016/j.tsc.2021.100864

29 Yıldırım, B. (2021). Preschool STEM activities: Preschool teachers' preparation and views. *Early Childhood Education Journal, 49*, 149–162. https://doi.org/10.1007/s10643-020-01056-2

Journal of Theoretical Educational Science, 17(1), 57-79, January 2024
Kuramsal Eğitimbilim Dergisi, 17(1), 57-79, Ocak 2024
[Online]: http://dergipark.org.tr/akukeg
DOI number: http://doi.org/10.30831/akukeg.1309078

The Effects of Out-of-School Learning Environments on STEM Education: Teachers' STEM Awareness and 21st-Century Skills*

Okul Dışı Ortamlarda STEM Eğitiminin Öğretmenlerin STEM Farkındalıklarına ve 21. Yüzyıl Becerilerine Etkisi

Mustafa ÇEVİK** **Büşra BAKİOĞLU***** **Zeynep TEMİZ******

Received: 2 June 2023 **Research Article** **Accepted:** 30 October 2023

ABSTRACT: This study was carried out within the scope of a nature education and science camp project. It aimed to investigate the effects of STEM activities conducted in out-of-school environments on teachers' STEM awareness and 21st-century skills. The research was designed according to the explanatory sequential pattern of the mixed method, and the quantitative and qualitative methods are the continuation of each other. Twenty-three science, elementary math, technology, and design teachers who worked at public middle schools participated in a six-day program that integrated STEM subjects with out-of-school learning. Seven different STEM-integrated out-of-school activities were implemented for six days in one of the central Anatolian regions in Turkey. STEM awareness and 21st-century skills scales were implemented in the study for data collection. A semi-structured interview was also conducted with teachers. The aim was for the teachers who participated in the project to integrate out-of-school and STEM education and then transfer this newly gained knowledge to their students. Both quantitative and qualitative findings support that teacher's acquisition of knowledge and skills throughout the science camp would support their students' craft, analytical thinking, motivation, problem-solving skills, and self-esteem. As such, out-of-learning environments should be used more in both STEM and 21st-century skills education.

Keywords: Out-of-school learning, middle school teacher training, STEM education, 21st century skills.

ÖZ: Bu çalışmanın amacı okul dışı ortamlarda gerçekleştirilen STEM etkinliklerinin öğretmenlerin STEM farkındalıklarına ve 21. yüzyıl becerilerine etkisini incelemektir. Araştırmada karma yöntemin açıklayıcı sıralı deseni benimsenmiştir. Devlet ortaokullarında çalışan 23 fen, matematik, bilişim, teknoloji ve tasarım öğretmeni, STEM konularını okul dışı ortamlarla bütünleştiren altı günlük bir programa katılmışlardır. Yedi farklı okul dışı ortamda kurgulanan STEM etkinlikleri, Türkiye'nin Orta Anadolu bölgesindeki bir ilde bir hafta boyunca yürütülmüştür. Araştırmada veri toplama amacıyla STEM farkındalık ve 21. yüzyıl becerileri ölçekleri kullanılmıştır. Ayrıca öğretmenlerle yarı-yapılandırılmış görüşmeler yapılmıştır. Araştırmaya katılan öğretmenlerin okul dışı ortamlara entegre edilen STEM eğitimleriyle edindikleri bilgileri öğrencilerine aktarmaları hedeflenmiştir. Çalışmanın bulguları, öğretmenlerin bilim kampı boyunca edindikleri bilgi ve becerileri, öğrencilerinin analitik düşünmelerini, motivasyonlarını, problem çözme becerilerini ve özgüvenlerini desteklemede kullanacaklarına işaret etmektedir. Bu nedenle hem STEM hem de 21. yüzyıl becerileri eğitiminde okul dışı ortamların daha fazla kullanılması önerilmektedir.

Anahtar kelimeler: okul dışı öğrenme ortamları, ortaokul öğretmen eğitimi, STEM eğitimi, 21. yüzyıl becerileri.

* This study was carried out within the scope of TUBITAK 4004 Nature Education and Science Schools Project and was presented as an oral presentation at the International Education Congress.

** Assoc. Prof. Dr., Karamanoğlu Mehmetbey University, Karaman, Türkiye, mustafacevik@kmu.edu.tr, https://orcid.org/0000-0001-5064-6983
*** *Corresponding Author*: Assoc. Prof. Dr., Karamanoğlu Mehmetbey University, Karaman, Türkiye, busrabakioglu@kmu.edu.tr, https://orcid.org/0000-0001-7997-1018
**** Assoc. Prof. Dr., Karamanoğlu Mehmetbey University, Karaman, Türkiye, ztemiz@kmu.edu.tr, https://orcid.org/0000-0002-4436-9737

Citation Information
Çevik, M., Bakioğlu, B., & Temiz, Z. (2024). The effects of out-of-school learning environments on STEM education: Teachers' STEM awareness and 21st-century skills. *Kuramsal Eğitimbilim Dergisi [Journal of Theoretical Educational Science]*, *17*(1), 57-79.

ISSN: 1308-1659

Formal education refers to structured, systematic, certificated, and generally state-supported learning that follows an organized curriculum (Ngaka et al., 2012). On the one hand, non-formal learning environments would be "real OSLE (OSLE)" such as museums, science centres, industrial institutions, zoos, and national parks. There may also be "digital/virtual OSLE" such as social media, educational websites, and web 2.0 tools (Karademir, 2018). There has been increasing attention on non-formal learning environments in formal education. Out-of-school education is not only effective for developing leisure-time interests and tuning in with the immediate environment, but it also helps students bring what they learn in the classroom into real life through direct observation, implementation, and experience. The out-of-school environment could be considered a laboratory in physical and social sciences, arts and crafts, health, and physical education (Smith, 2017).

Learning that takes place in non-formal learning environments is not completely independent of the formal teaching offered in the classroom. Rather, a link could easily be established, as most institutions propose learning goals and employ educated staff. However, these easily accessible learning environments are often visited in an unstructured, recreational way. When detailed aims, concepts, or competencies are explicated, such informal settings may offer crucial benefits (Randler et al., 2012). In OSLE, the education curriculum and methodology are flexible. Still, learning takes place in an intentional and organized way, not just by chance. This flexibility empowers students to place their needs and interests at the centre of education. While formal education pays little attention to skill development, out-of-school education emphasizes skills and the development of attitudes such as tolerance. Students' needs are better met in OSLE, which allows them to know themselves and the world more deeply (Grajcevci & Shala, 2016).

According to the STEM Education Turkey Report, STEM education should not be limited to in-class activities; rather, it should be supported with OSLE (Akgündüz & Ertepınar, 2015). There are several studies on the learning outcomes of STEM activities performed in OSLE (Bakioğlu & Çevik, 2021; Baran et al., 2016) and attitudes toward STEM (Cooper & Heaverlo, 2013; Timur et al., 2020; Young et al., 2017). Yet, to the best of our knowledge, there are no known studies focusing on the effects of STEM activities carried out in out-of-school environments on teachers' STEM awareness and 21st-century skills. This research is expected to contribute to the literature on the effects of STEM activities in OSLE. It is a well-known fact that science, technology, engineering, and math took their actual origin from nature; therefore, it is preferable for them to be taught in natural settings. All materials, phenomena, procedures, and problems to be solved could easily be found in nature. Similar to nature, additional out-of-school environments offer enriched educational opportunities for both STEM awareness and 21st-century skills.

This research was carried out for one week with 23 teachers from four different branches in seven different out-of-school environments within the scope of the project called "STEM is All Around," supported by the TÜBİTAK 4004 project. The main aim was to find and describe useful teaching methods in STEM education (SE), including out-of-school education, for promoting 21st-century skills of middle school teachers. This study is guided by the following research main question and sub-question:

1. How did the STEM activities implemented within the project contribute to the teachers?

1.1. Does STEM education conducted in out-of-school learning environments influence teachers' STEM awareness?

1.2. Does STEM education conducted in out-of-school learning environments influence teachers' 21st-century skills?

1.3. What are the participants' opinions on how STEM education contributes to them in out-of-school learning environments?

Theoretical Framework and Literature Review

It has been revealed that out-of-school learning environments increase the social skills of individuals (Panizzon & Gordon, 2003), and such environments enable teachers to learn new ideas and information from each other and increase and improve their teaching strategies and skills (Chin, 2004). STEM education, whose starting point is the philosophy of holism (Simmons, 2021). The multidisciplinary nature of it can be attributed to the influence of Gestalt theory. The Gestalt Approach helps create new ways of discovering, experimenting, or creating educational content, where the whole is more than the parts that make it up and helps capture the development of learning environments that are more connected to the students' universe, values subjectivity, uniqueness, responsibility and experience (Greenwood, 2020). This philosophical doctrine focuses on the perception of existence in the face of a holistic environment; It coincides with the transdisciplinary logic of the STEM approach. From this point of view, STEM is more than just an approach applied in the classroom. STEM education, which has been popular in recent years, can also be applied in out-of-school learning environments such as summer camps, museum education, and so on. One of the most important reasons for this is that STEM education takes place in out-of-school learning environments is that it offers better foundation for the curriculum content in the USA (Özçelik & Akgündüz, 2018). The goals of STEM education include increasing the knowledge, skills, attitudes, and career awareness of individuals towards STEM as well as the development of 21st-century skills. Out-of-school learning environments contribute to the development of these goals (Bozkurt et al., 2009). It is important that teachers who are expected to transfer STEM goals to their students have high STEM awareness, as teachers with high STEM awareness will provide STEM education more effectively.

Today's world is changing at an increasing speed, requiring new and shifting knowledge and skills. The 21st-century skills include traditional curricular content such as reading, writing, mathematics, history, and science, as well as writing, interpretation, and synthesis. The Organization for Economic Co-operation and Development (OECD) defines 21st-century skills as cognitive skills, intrapersonal skills, and technical skills (OECD, 2019). Cognitive skills combine thinking, learning, remembering, and reasoning to transform problem-solving and critical thinking. Intrapersonal skills encompass a group of self-communication skills to manage emotions and reach executive functioning. Technical skills involve performing practical tasks that naturally require information fluency in using certain tools and technologies (Ananiadou & Claro, 2009).

The American National Science Teachers Association (NSTA) emphasizes that 21st-century skills have a natural and strong relationship with science education. It is reported that science applications that specifically address the nature of science can provide a rich context for the development of many 21st-century skills, such as critical thinking, problem-solving, and information literacy (NSTA, 2011). Students can be provided with authentic learning environments to develop 21st-century skills (Larson & Miller, 2011). Besides, allowing students to scrutinize curricular tasks to construct their knowledge triggers the development of 21st-century skills (Bybee, 2009).

Similarly, STEM education is the core focus of many countries around the world. Each nation is converting its educational systems to be competitive in the 21st century. However, the situation is not promising in Turkey, as the majority of Turkish students lack access to STEM experiences that meet international standards in STEM disciplines. One of the most important reasons for this is that there are not enough teachers trained in the STEM field (Çorlu et al., 2014). Consequently, STEM activities carried out in an out-of-school environment will also help teachers to improve themselves, who will then raise qualified individuals (Özbilen, 2018).

OSLE provides students with unique and enriched learning opportunities that are not available during the regular school program. This enrichment increases with relevant course content and STEM practices that are not possible in a traditional classroom setting (Peterson & Fix 2007). STEM education aims to increase students' knowledge, skills, and attitudes towards STEM, as well as improve their 21st-century skills. As such, non-formal learning environments play a crucial role in contributing to STEM education goals (Bozkurt Altan et al., 2019). The STEM education program integrated into non-formal learning environments provides cooperative, practical learning opportunities that enable students to find solutions for their daily life problems (Schnittka et al., 2010). Activities that take place outside of school increase students' interest in the subject matter (Bogner & Wiseman, 2004). Teachers shoulder great responsibility to establish a bridge between STEM education and non-formal learning environments to enhance student development. As such, this study might be useful in filling the mentioned gap in STEM education in out-of-school learning environments.

Method

This research employed a mixed-methods approach to examine the effects of STEM activities conducted in OSLE on teachers' STEM awareness and their development of 21st-century skills. Yıldırım and Şimşek (2013) stated that one of the most important reasons for the use of mixed designs in research is that the world around us is complex and comprehensive. The basic premise for choosing the mixed method is to add depth and breadth to the study as the holistic picture from meanings obtained from interviews or observation to the generalization sample from a population obtained from surveys. A mixed-methods study provides stronger inference than using quantitative and qualitative approaches alone (Wasti et al., 2022). In the current research, explanatory sequential design was preferred among mixed designs as it is effective in explaining qualitatively the quantitative results that may occur unexpectedly or unusually (Morse, 1991; Plano-Clark & Creswell, 2008).

Study Group

The science camp was held in September of the 2021-2022 academic year. The study group was determined using criterion sampling, and the participants were selected according to criteria determined by experts. Their willingness to participate in the project, lack of project experience, and having a master's degree were effective in selecting the participants. Twenty-three science, elementary math, technology, and design teachers who worked at public middle schools were chosen from 154 online applications from eight different cities.

Table 1

Demographic Characteristics of Participating Teachers

		f	%
Gender	Female	17	73.9
	Male	6	26.1
Major	Science	14	60.8
	Elementary Maths	5	21.8
	Information Technology	3	13.1
	Technology & Design	1	4.3
Type of school	Public	22	95.6
	Private	1	4.4
Term of office	1-5	5	21.7
	6-10	7	30.4
	11-15	9	39.2
	16 and over	2	8.7
Project experience	Yes	5	21.7
	No	18	78.3

Table 1 shows that 17 of the teachers in this study are female, and six are male. Fourteen of the teachers are from science, 5 of them are from primary school mathematics, 3 of them are from information technologies, and 1 of them is from the technology design branch. Twenty-two of the teachers work in public schools, and one of them works in a private school. Five of the teachers have 1-5 years, 7 have 6-10 years, 9 have 11-15 years, and 2 of them have 16 years or more years of work experience. Five of the teachers have project experience, and 18 do not have any project experience in the current research.

Data Collection Instruments

"STEM Awareness Scale" (15 items) and "Multidimensional 21st Century Skills Scale" (41 items) were used for quantitative data collection. Reliability analyses of the scales were conducted. The data were analyzed using the SPSS-24 program based on normality tests. In line with the findings obtained from the quantitative data, the

qualitative stage was initiated. Participating teachers were interviewed in person using a semi-structured interview form. Then, the audio recordings and transcriptions of the interviews were analyzed using content analysis.

Quantitative Data Collection Instruments

STEM Awareness Scale

The scale developed by Çevik (2017) was used to determine teachers' awareness of STEM. This scale was prepared in a 5-point Likert style and consisted of a total of 15 items and three factors. The first factor measures the effect of teachers' STEM awareness on students, the second sub-dimension measures the effect on the lesson, and the third sub-dimension measures the effect on the teacher. It was confirmed by CFA that the scale consisted of 3 sub-dimensions. The Cronbach's Alpha coefficient for the overall scale is .82., and the reliability coefficients of the factors are .81, .71, and .70. The test-retest method was used to validate the scale. In the first stage of this method, the standard deviation and mean of the scale were determined as 0.52 and 3.95, respectively. In the second stage, they were determined as 0.53 and 3.91, respectively. Pearson Correlation Coefficient was found to be significant at the level of p=.001 and r= .615.

Multidimensional 21st Century Skills Scale

The second scale used in the research was developed by Çevik and Şentürk (2019). The scale is in a 5-point Likert style and consists of 41 items, with five factors: "Information and Technology Literacy Skills," "Critical Thinking and Problem-Solving Skills," "Entrepreneurship and Innovation Skills," "Social Responsibility and Leadership Skills" and "Career Awareness". The alpha reliability coefficient of the scale is .86 for the overall scale, and for the sub-dimensions of the scale, it is .84, .79, .76, .73, and .75, respectively. The agreement results obtained with the DFA analysis regarding the reliability of the scale are (χ^2/sd = 2014.17/774; GFI = .90; AGFI = .85; RMSEA = .050; CFI = .95; NFI = .91 and SRMR = .058), which are acceptable for the application. The Cronbach's Alpha reliability coefficient of the scale was determined as .92 for this study.

Qualitative Data Collection

Semi-structured Interview Protocol

In line with the analysis of quantitative data, a semi-structured interview protocol was used to determine the reasons for the changes identified in the participants. It was developed to reveal the main reasons for the changes in 21st-century skills and a STEM awareness in depth. Validity and reliability concepts are against the quantitative paradigm; credibility and trustworthiness are generally used in qualitative research (Krefting, 1991).

In qualitative research, the researcher should ensure that the findings are drawn from the data, not the interpretation of the researcher. Confirmability helps the researcher to reduce bias originating from the researcher's own perspective. Several strategies have been implemented to ensure the credibility of the study. First of all, researchers developed rapport with participants during the science camp, which led to

more objective responses in the interviews. Secondly, an audit trail is an effective way to offer a clear description of the research process (Shenton, 2004). The current study employed an audit trail to increase confirmability. The process and results were reviewed by another STEM field expert besides the researchers. Additional expert was asked to compare the same themes with conceptual categories so that no themes were left outside. Thirdly, intercoder agreement was determined by using the formula of Miles and Huberman (1994) Reliability= consensus/ (agreement+ disagreement) *100. According to Miles and Huberman (1994), agreement among coders is expected to be at least 80%. The intercoder agreement was accepted as satisfactory as it reached 84% (83/83+15). In order to ensure confidentiality, the participating teachers were named T1 to T23 (T stands for the teacher, and the accompanying number stands for the number of participants), and the most representative quotations were selected. Initial findings were presented at international conferences for peer scrutiny. Finally, the sample selection was made for transferability, as the characteristics of the participants and the research field explained (Sharts-Hopko, 2002). The most frequently used sampling methods in qualitative research are purposeful, random, and voluntary sampling. How the participants were selected in the study was stated objectively.

The interview questions were formed as follows.

1. Which activity do you think contributed the most to you during the science camp? Explain with reasons.

2. Which activity do you think contributed the least to you during the science camp? Explain with reasons.

3. In which aspect do you think the science camp and STEM education developed you?

4. Which STEM discipline would you say had the most impact on you in science camp?

5. In which aspects do you think your awareness increased in the science camp?

Activities Procedure

The science camp lasted for a total of 6 days in OSLE. These environments included Değle Ruin, a coding laboratory, a playground, a science centre, a museum, a hospital, and a natural park. Legal permissions and ethical documents were obtained from the relevant authorities. Within the scope of the science camp, STEM activities included research and inquiry-based science activities within the framework of engineering, biomimicry, design, coding, environmental awareness, and renewable energy themes. Measurement and evaluation studies were also organized in order to determine the teachers' level of basic knowledge and skills they will acquire in the science camp. Additionally, warming drama games were held for participants to get to know each other. In this study, all activities are genuine and developed by researchers. First, out-of-school environments were determined. Then, the literature was investigated for appropriate activities. For the following step, activities were sent to prospective trainers, and based on their feedback, necessary adjustments were made to finalize activities.

First Day

The participants arrived at the drama class of Karamanoğlu Mehmetbey University Faculty of Education in the morning. They participated in four different warming drama activities, which were prepared to help them get to know each other and express themselves. After the activities, the participants learned each other's names, fields of study, professional experiences, and the provinces they came from to participate in the project. The session focused on planning STEM education in out-of-school settings, providing examples of how STEM education can be integrated into the curriculum and how teachers can expand their teaching beyond the classroom through collaborative, hands-on, and project-based learning activities. STEM activity worksheets and lesson plans were prepared and presented to show how to integrate STEM disciplines into out-of-school activities. To provide a real-life experience, a visit to an apple orchard was organized.

Second Day

Remodeling of Değle Ruins (Karadağ): Değle Ruins is a settlement from the Ancient Eastern Roman period, located 50 km from Karaman-Center. A local resident welcomed the participants at Değle Ruins and provided information about the ruins. An expert in the field of STEM gave a presentation by projecting onto ancient walls, and then the participants took measurements of ancient buildings with laser meters, using both standard and non-standard units. Finally, the group visited the village's primary school to create virtual and real models in a 1:64 ratio based on their measurements.

Third Day

Smart Lighting System Design in the Coding Laboratory: The coding laboratory at the Youth Centre offers an excellent premise for out-of-school learning. The TÜBİTAK Experimental Laboratory has outstanding technical and technological equipment for coding experiences. In addition to the coding training, the participants also examined the laboratories, materials, and social facilities in the youth centre. They set up a lighting system and coded it using the Tinkercad Program. They also learned how they can benefit from or contribute to Youth Centres in their cities. Designing a Toy using STEM in the Parks: On the third day afternoon, the participants made extensive observations of the Karamanoğlu Mehmetbey University Practice Kindergarten Playground materials. They examined the working principles of the materials in the park. Then, the participants designed a toy consisting of simple machines within the scope of STEM.

Fourth Day

Let's Rebuild the Solar System: Konya Science Centre is one of the largest science centres in Turkey. Science centres are attractive not only because of their content but also because of their architecture, green spaces, and diverse usage purposes. TÜBİTAK aims to spread science culture in our society with the help of the science centre. Thus, they are well-suited to the philosophy of out-of-school environments. A planetarium trip was carried out in the centre, and the participants then created a solar system by reducing the sizes of the planets and the distances between them. After the

participants examined each other's systems and finished their presentations, they examined the materials, devices, and automation systems in interactive laboratories.

Fifth Day

Design a Simple Machine using STEM Education in Museums: The materials were presented chronologically to the participants by an expert in the museum. The participants gained technical knowledge by examining the simple machines and equipment and received information about their working principles. In particular, the relationship between the number of gears and the force applied by coffee and flour grinders was explained. Then, the participants developed their own simple machines within the scope of the STEM approach.

Muscle Measurement Device Design in Muscle Research Laboratory: The second activity of the fifth day was carried out at the Muscle and Athlete Exercise Evaluation Laboratory in Karamanoğlu Mehmetbey University Medical Faculty Hospital. It was explained how to benefit from health institutions as an out-of-school learning environment and how to protect against the risk of infection and transmission. Participants' ideas on how to benefit from health institutions were taken, and muscle strength was measured in one participant. Then, the participants designed their own muscle-measuring devices, and all designed devices were tested.

Sixth Day

STEM in Nature - Design Transportation Vehicle: Theoretical knowledge about biomimicry was provided to the participants one day in advance. The activity worksheet, which they used during their observations in nature, was distributed to the participants. Information was obtained from the authorities about the nature park and the creatures found in Yer Köprü Waterfall. At the end of their trip, the participants created appropriate drawings for the vehicle they wanted to design by taking advantage of the characteristics of the living things in the nature park. They then designed the transportation vehicles with the materials provided to them. All groups presented the tools they designed, and an evaluation was conducted.

Data Analysis

The data analyses were conducted using the SPSS 24 software package. As the study group comprised less than 50 people, the Shapiro-Wilk test was used (Büyüköztürk et al., 2011). However, the normality test did not indicate a normal distribution. Besides, the skewness and kurtosis values were outside the range of +1.5/-1.5 (Tabachnick & Fidell, 2007). Therefore, changes in the group from the pre-test to the post-experiment were analyzed using the Wilcoxon Signed Rank Test, one of the non-parametric tests (Privitera, 2015; Wilcox, 2012). Qualitative data was analyzed using content analysis, as explained above.

Ethical Procedures

Ethical approval for the research was obtained from the Karamanoğlu Mehmetbey University Ethics Committee with the number 07-2022/187 and date 09.11.2022. After obtaining ethical approval, the participants were asked to sign a consent form indicating that they were voluntarily participating in the study.

Results

In this section, answers to the sub-questions under the main question: "In which direction did the STEM activities realized within the scope of the project contribute to the teachers?"

The following findings are in line with the first research sub-question, "Does STEM education conducted in OSLE influence teachers' STEM awareness?" The differences between the pre-test and post-test scores of the participants on the STEM Awareness Scale, both in general and its sub-dimensions, as a result of non-parametric analyses are given in Table 2.

Table 2

Wilcoxon Signed-Rank Test Results for The Participant Pre-Test and Post-Test STEM Awareness Scale

	Pre-test – Post-test	*n*	Mean Rank	Sum of Ranks	*Z*	*p*
Effect of Student	Negative Ranks	1	2.50	2.50	-3.29	.00*
	Positive Ranks	12	7.38	88.50		
	Ties	10				
Effect of Course	Negative Ranks	11	10.64	117.00	-3.64	.71
	Positive Ranks	11	12.36	136.00		
	Ties	1				
Effect of Teacher	Negative Ranks	7	9.50	66.50	-.107	.24
	Positive Ranks	12	10.29	123.50		
	Ties	4				
General	Negative Ranks	9	8.83	79.50	-1.67	.12
	Positive Ranks	13	13.35	173.50		
	Ties	1				

Table 2 shows that after attending the OSLE STEM activities (museum, science centre, playground, ruins, youth centre, research hospital), participants' STEM awareness increased. A significant increase was observed in favour of the "effect of student" sub-dimension (z=-3.29, p<.05). However, this difference was not significant across the scale and in other sub-dimensions. Participants' 21st-century skills were measured with the Multidimensional 21st Century Skills Scale. Result of non-parametric analyses are given in Table 3.

The following findings are in line with the second research sub-question, "Does STEM education conduct in OSLE influence teachers' 21st-century skills?"

Table 3

Wilcoxon Signed-Rank Test Results for The Participant Pre-Test and Post-Test Multidimensional 21st Century Skills Scale

	Pre-test – Post-test	*n*	Mean Rank	Sum of Ranks	*z*	*p*
Knowledge and technology literacy skills	Negative Ranks	4	11.88	47.50	-2.51	.01*
	Positive Ranks	18	11.42	205.50		
	Ties	1				
Critical thinking and problem-solving skills	Negative Ranks	14	9.75	136.50	-1.12	.23
	Positive Ranks	6	12.25	73.50		
	Ties	3				
Entrepreneurship and innovation skills	Negative Ranks	5	6.60	33.00	-2.69	.00*
	Positive Ranks	15	11.80	177.00		
	Ties	3				
Social responsibility and leadership skills	Negative Ranks	8	12.06	96.50	-.98	.32
	Positive Ranks	14	11.18	156.50		
	Ties	1				
Career consciousness	Negative Ranks	6	10.67	64.00	-.94	.34
	Positive Ranks	12	8.92	107.00		
	Ties	5				

Statistically significant differences were observed on the "Multidimensional 21st Century Skills Scale". The difference was observed in sub-dimensions of "Knowledge and Technology Literacy Skills" (z=-2.51, p<.01) and "Entrepreneurship and Innovation Skills" from the (z=-2.69, p<.05) as shown in Table 3. The difference in scores in favour of positive ranks in these sub-dimensions shows that the activities carried out within the scope of the science camp have a significant effect on the Multidimensional 21st Century skills of teachers, specifically Knowledge and Technology Literacy Skills and Entrepreneurship and Innovation Skills. No statistically significant difference was found in the sub-dimensions of Critical Thinking and Problem-Solving Skills, Social

Responsibility and Leadership Skills, and Career Consciousness In order to investigate the underlying reasons for the quantitative findings, interviews were conducted with the participants to answer the third problem (What are the participants' opinions on how STEM education contributes to them in out-of-school learning environments?) of the research. The interviews were coded openly, and five themes emerged: Most effective and least effective activities, efficacy in teaching practice, creating awareness, and contributing to STEM disciplines. These themes are represented in Table 4, Table 5, Table 6, and Table 7 below.

Table 4

The Activities That the Participants Found the Most and Least Effective to Them in the Research

Main Themes	Answers	Codes
Ruins	It contributed a lot (19)	* Technology use (11)
		*Exploring different programs (9)
		*Bringing different disciplines together (5)
	It contributed little (4)	*Insufficient internet (3)
		* Transport (1)
Coding Workshop	It contributed a lot (15)	*Using technology (10)
		*Exploring different programs (10)
	It contributed little (8)	*Low readiness (6)
		*Mixed (4)
Science Centre	It contributed a lot (20)	*Using technology (10)
		*Using different materials (9)
		*Creativity development (8)
		*Dexterity (4)
	It contributed little (2)	* Inadequate in self-education (1)
	No response (1)	
Children play area	It contributed a lot (18)	* Opportunity to develop yourself (10)
		*Imagination development (10)
		*Development of dexterity (8)
		*Ensuring seeing (2)
		*Using technology (1)
	It contributed little (3)	*STEM was not complete (1)
	No response (2)	*Materials were insufficient (1)

Muscle Lab	It contributed a lot (15)	*Using technology (9) *Development of dexterity (5)
	It contributed little (5)	*Use of simple materials (5) * STEM was not complete (4)
	No response (3)	
Museum	It contributed a lot (20)	*Creativity (12) *The importance of group work (4) *Integrating different disciplines (4) *Craft development (4)
	It contributed little (3)	*The setting was not suitable for STEM (2) * More complex problem could be solved (1)
Nature Park	It contributed a lot (18)	*Creative feature (8) *Dexterity (5) *Imagination (4)
	It contributed little (3)	*Using technology (3)
	No response (2)	*Environment compelling (2)

The activities that teachers find most effective, ruins, science centre, and museum, respectively, are shown in Table 4. Teachers benefited from these activities mostly in terms of using technology (21), exploring different programs (19), creativity (20), integrating different disciplines (9), and developing fine motor skills (8). T20 said, "I learned how to construct simple machines with the STEM activity in the museum, my creativity improved," while T5 asserted that "STEM education and modelling work in archaeological sites impressed me a lot, I learned a lot of technology" again T8 stated "Of the activities, the science centre impressed me the most and inspired me how the disciplines came together. It has opened up new horizons for the future." Several ancient civilizations have been founded and collapsed in Anatolia. As such, around 150 ruins can be found all over Turkey, but there is a dearth of studies focusing on using those ruins for educational purposes. Besides, all provinces and surrounding towns have lots of museums. Although museum education has been popular for a couple of decades, it is also rare in Turkey. Teachers mostly expressed that the activities held in ruins and museums improve their creativity and use of technology skills. These findings were also supported by STEM awareness and 21st-century skills scale findings.

Table 5

Opinions of the Participants of the Research on Which Direction They Develop More within the Scope of the Teaching Profession

Main Themes	Answers	Codes
Student training	It was helpful (22)	*Technology use (8)
		* Doing different activities (6)
		* Variety of out of school environments (6)
	No response (1)	* STEM approach (5)
Professional development	Has a positive contribution (23)	*Using technology (10)
		*Exploring different programs (8)
		*Current topics (6)
		*Imagination (4)
		*Crafts (4)
Course	It was helpful (21)	*Using technology (9)
		*Using different materials (9)
		*Content enrichment (6)
		*Different approaches (4)
		*Preparing lesson plan (4)
	No response (2)	

Teachers stated that the science camp contributed to their professional development and teaching as well as beneficial to students (Table 5). They reported that their acquisitions in using technology (8), doing different activities (6), experiencing a variety of out-of-school environments (6), and implementing a STEM approach (5) will contribute to their teaching under the theme of raising students. They further stated that acquisitions such as using technology (10), exploring different programs (8) and current topics (6), imagination (4), and hands-on trainig (4) will improve their professional development. Again, they stated that acquisitions such as using technology (9), using different materials (9), enriching the content (6), different approaches (4), and preparing a lesson plan (4) would positively contribute to their teaching under the course theme. T1 exemplified the issue as “The technological tools and materials we use in the science camp will allow us to use them with our students,” T3 said, “We learned that there are different out-of-school environments in the science camp. I will conduct similar activities with my students when I return to my school.” Again, T10 stated, “I will use the materials such as the lesson plans and worksheets that we used in the activities in the science camp,” and T2 expressed his improvements by reporting, “The products we developed in the science camp increased my hand skills and imagination. I feel like a different teacher.” The activities were all practical, and teachers engaged in group work to design a product; they asserted that their learning was efficient, permanent, and transferrable to their teaching. Teachers expressed that their acquisition of knowledge and skills throughout the science camp would support their students’ craft, analytical thinking, motivation, problem-solving skills, and self-esteem, which also appeared in STEM awareness scale findings.

Table 6

Opinions of the Participants on Which Discipline that Compose STEM Has More Influence in the Science Camp

Main Themes	Answers	Codes
STEM	Quite effective (23)	Science (7), Technology (20), Engineering (15), Mathematics (5)

Participating teachers stated that the science camp is most effective in STEM disciplines, specifically Technology (20), Engineering (15), Science (7), and Mathematics (5) (Table 6). T7 summarized her gains in STEM disciplines during the science camp, stating, "We used technology and design disciplines more in the science camp, which made me feel its effect more." Similarly, T5 said, "In general, technology and programs affected us more in the science camp," and T1 said, "I think we used science more; it was at the core of all activities."

Table 7

Opinions of the Participants on Which Direction Their Awareness Increased Most with the Science Camp

Main Themes	Answers	Codes
STEM	Increased (22)	What is STEM (12) Gathering disciplines together (7) Designed based approach (6)
	No response (1)	
Out of School	Increased (22)	Out of school (16) Different activity in out of school (10)
	No response (1)	
Team Work	Increased (21)	Importance of working together (14) Group work (9)
	No response (2)	
Imagination	Increased (20)	Creativity (10) Designing by imagination (10) Multidimensional thinking (5)
	No response (3)	

Table 7 represents an increase in teacher awareness in four main areas during the science camp. Teachers stated that they now comprehend STEM better (12), can bring disciplines together (7), learned a design-based approach under the STEM theme (6),

and became aware of out-of-school environments (16) and different out-of-class activities (10). Besides, they understood the importance of working together (14) and the significance of doing activities with the group (9). Finally, under the main theme of Imagination, teachers stated that they became more aware of the skills of creativity (10), designing by imagination (10), and multidimensional thinking (5). T4 said, "I realized what STEM is, especially with the science camp." T17 added, "I realized how important it is to have the imagination to make design." T20 said, "I became aware of out-of-school environments during the science camp."

Participating teachers came from all over the country. Although their outdoor environments, facilities, and opportunities in their hometown would be different, they basically experienced and learned how to integrate STEM and 21st-century skills in an out-of-school environment. Out-of-school experiences, especially the mentioned subject, are not common in Turkey. Therefore, teachers emphasized that the knowledge and skills that they gained through the science camp would enrich their students learning.

Discussion

In recent years, it has been accepted that the STEM approach has a positive effect on the development of individuals' scientific process skills in activities such as nature education, science festivals, and science schools. The STEM approach also has a positive effect on students' perspectives on science, the nature of science, and their attitudes towards science and science laboratory courses (Balım et al., 2013; Çelik, 2012; Çevik & Abdioğlu, 2018; Markowitz, 2004). This current study aims to identify the effective use of out-of-school environments for STEM education.

The findings of this study are consistent with previous findings that out-of-school STEM activities increase teachers' STEM awareness. Teachers' STEM awareness is statistically significant in favour of the "effect on students" sub-dimension. Yet, this difference was not significant in the course and teachers' sub-dimensions. Within the scope of the research, it can be concluded that the activities carried out significantly increased teachers' STEM awareness. Quantitative findings were supported by qualitative results that teachers reported that their STEM awareness particularly increased in students, lessons, and teaching content (Aslan-Tutak et al., 2017; Karısan et al., 2019). The effect of STEM activities carried out in OSLE may be higher as OSLE provides enthusiasm for learning, hands-on experiences, richness of stimulus, and a real-world context (Akaygün et al., 2015).

After attending the STEM activities in OSLE, the 21st-century skills of the participating teachers increased, and this was observed more in the Entrepreneurship and Innovation Skills and Critical Thinking and Problem-Solving Skills sub-dimensions. No significant difference was found in Critical Thinking and Problem-Solving Skills, Social Responsibility and Leadership Skills, and Career Consciousness sub-dimensions. These findings are also supported by the findings of the qualitative aspect of study. Teachers reported that they became more enthusiastic about solving problems and interested in developing solutions to make life easier. They also stated that they approached people with different personalities more positively, did not avoid solving the difficult problems they encountered, and did not accept the information they learned as it was but accepted it after evaluating it with a critical approach. These

findings support the hypothesis of the study as the STEM approach provides interdisciplinary integration and contributes to the development of individuals' innovative problem-solving skills (Roberts, 2012; Schnittka, et al., 2010; Şahin, et al., 2014). Studies have revealed that STEM education helps teachers improve their 21st-century skills (Cunningham & Kelly, 2017; Sullivan, 2008). This study is also consistent with the literature in this aspect since participating teachers contribute to 21st-century skills as they are involved in hands-on activities in OSLE, which offer more stimulus, materials, and phenomena than they can see inside the classroom. Therefore, teachers also reported that they sometimes faced some difficulties during the STEM activities. They stated that they were limited by factors such as lack of internet and some infrastructure deficiencies. Similarly, Thomas (2010), Koosimile (2004), Oriaon et al. (1997), and Tatar and Bağrıyanık (2012) also reported that teachers face similar challenges in out-of-school settings.

Qualitative findings of the study revealed that participating teachers mostly benefited from the historic site, science centre, and museum. Also, their use of technology, discovery of different programs, creativity, integration of different disciplines, and hand skills improved during science camp. Science centers and museums help individuals understand the nature of science by developing their questioning skills (Kubat, 2018). Science education in science museums provides permanent learning and supports classroom learning (Martin, 2004). Ruins, science centres, and museums offer several interesting stimuli for STEM education; this is why teachers stated that they included these out-of-school environments. They used technology more on these sites as they made several measurements with different tools and applied their measurement to both virtual and hands-on designs.

Teachers asserted that the science camp contributed to the themes of being more beneficial to the students in the context of professional competence, professional development, and teaching. Under the main theme of raising students, teachers believe that their acquisitions, such as using technology, doing different activities, incorporating diversity of out-of-school environments, and using the STEM approach, will contribute to educating their students. Under the theme of professional development, they stated that their use of technology, discovering different programs, current issues, imagination, and hand skills improve professionally. Under the lesson theme, they stated that the acquisitions, such as the use of technology, using different materials, enriching the content, using different approaches, and preparing a lesson plan, will contribute positively to the teaching they give in their schools. The findings support Bakırcı and Kutlu's (2018) findings that the STEM approach will increase students' interest and motivation towards the lesson, enable them to think multi-dimensionally, develop their research-inquiry and creativity skills, design products suitable for the problem situation, learn by embodying the subjects, and develop their scientific process skills.

The teachers reported they were influenced by STEM disciplines in descending order: technology, engineering, science, and mathematics. Teachers are believed to be significantly influenced by technological tools and engineering, as they see technological tools (laser meter, muscle measuring device, wireless projection, etc.) that they have never seen before in these environments and make designs like an engineer inspired by these tools. Moreover, compared to other disciplines, technology, and engineering are more popular in media; thus, teachers might be more focused on these

disciplines. Similarly, as stated in the literature (Sungur Gül & Marulcu, 2014; Wang 2012), the use of technological tools and equipment by teachers motivates them more to STEM activities, especially in engineering fields.

Teachers' awareness increased in four main areas during the science camp. Teachers stated that they comprehend STEM better, bring disciplines together, and learn a design-based approach under the STEM themes. They also became aware of the out-of-school environments and different out-of-class activities. Besides, they understood the importance of working together and doing activities with the group. Finally, under the main theme of Imagination, teachers stated that they became more aware of the skills of creativity, designing by imagination, and multidimensional thinking. Teachers comprehend how to do STEM activities in out-of-school environments while they come together to design and create a product as a group. It is emphasized that critical thinking, problem-solving skills, being creative, and working collaboratively are among the 21st century skills (Akgündüz et al. 2015).

STEM education in OSLE enables people to take responsibility through teamwork, experience communication with colleagues, and observe how science works by doing their own experiments (The Parliamentary Office of Science and Technology [POST], 2011). To sum up, both quantitative and qualitative findings of the research revealed that STEM was an ambiguous concept for teachers before they attended the project. With the help of the theoretical and practical activities during the project, teachers gained a thorough understanding of the STEM approach. Consequently, teachers' STEM awareness and 21st-century skills improved.

Conclusion

Participating teachers' knowledge, experiences, and skills in applying STEM and 21st-century skills in out-of-school environments seemed to be limited due to the lack of sufficient pre-service and in-service training regardless of their major. Therefore, such training should be improved in order to provide better experiences for middle school students. One way to improve middle school teachers' out-of-school environment practices concentrated on STEM and 21st Century Skills would be to provide them with hands-on activities, such as those presented in this study. Activity books, online activity pools, and more science camps enable teachers to use out-of-school environments more effectively.

Implications

STEM activities in OSLE increase people's awareness of STEM and 21st-century skills. Therefore, it is recommended that educators organize more STEM activities in these environments. Enriched activities should be planned so that people can benefit more from STEM activities in OSLE. STEM activities that take place in OSLE should be carried out collaboratively as much as possible. It is observed in this study that teachers were not aware of the usefulness of ruins, museums, and youth centres, which are readily available in all the cities in Turkey.

Acknowledgements

This research covers a part of the project numbered 122B774 supported by TÜBİTAK 4004 Nature Education and Science Schools

Statement of Responsibility

Büşra Bakioğlu and Zeynep Temiz are responsible for the introduction, discussion and conclusion sections. Mustafa Çevik is responsible for the method, findings and data analysis section. All authors participated in writing and critical review.

Conflicts of Interest

The authors have no competing interests to declare that are relevant to the content of this article.

Author Bios:

Mustafa Çevik is an associate professor at the Faculty of Education, Karamanoğlu Mehmetbey University. His research interests are, science education, STEM education, artificial neural network, 21st century skills.

Büşra Bakioğlu is an associate professor at the Faculty of Education, Karamanoğlu Mehmetbey University. Her research interests are, science education, out of school environment, STEM education.

Zeynep Temiz is an associate professor at the Faculty of Education, Karamanoğlu Mehmetbey University. Her research interests are, early childhood education, STEAM education, out of school environment.

References

Akaygün, S., Aslan-Tutak, F., Bayazıt, N., Demir, K., & Kesner, J. E. (2015). *Kısaca FeTeMM eğitimi: Öğretmenler ve öğrencileri için iki günlük çalıştay.* 2. International Conference on New Trends in Education, İstanbul, Türkiye.

Akgündüz, D., & Ertepınar, E. (Eds.). (2015). *STEM eğitimi Türkiye raporu.* Scala Basım Yayım Tan. San. Tic. Ltd. Şti.

Akgündüz, D., Aydeniz, M., Çakmakçı, G., Çavaş, B., Çorlu, M., Öner, T., & Özdemir, S. (2015). *STEM eğitimi Türkiye raporu: Günümüz modası mı yoksa gereksinim mi?* İstanbul Aydın Üniversitesi STEM Merkezi.

Ananiadou, K., & Claro, M. (2009). *21St century skills and competences for new millennium learners in OECD countries.* OECD education working papers, no. 41. OECD Publishing (NJ1).

Aslan-Tutak, F., Akaygün, S., & Teksezen, S. (2017). Collaboratively learning to teach STEM: Change in participating pre-service teachers' awareness of STEM. *Hacettepe University Journal of Education, 32*(4), 794-816.

Bakırcı, H., & Kutlu, E. (2018). Fen bilimleri öğretmenlerinin FeTeMM yaklaşımı hakkındaki görüşlerinin belirlenmesi. *Türk Bilgisayar ve Matematik Eğitimi Dergisi, 9*(2), 367-389.

Bakioğlu, B., & Çevik, M. (Eds). (2021). *Okul dışı ortamlarda STEM eğitimi*. Nobel Akademi Yayıncılık.

Balım, A. G., Çeliker, H. D., Türkoğuz, S., & Kaçar, S. (2013). The effect of reflections of science on nature project on students' science process skills. *Journal of Research in Education and Teaching, 2*(1), 149-157.

Baran, E., Bilici, S. C., Mesutoglu, C., & Ocak, C. (2016). Moving STEM beyond schools: Students'perceptions about an out-of-school STEM education program. *International Journal of Education in Mathematics, Science and Technology, 4*(1), 9–19.

Bogner, F. X., & Wiseman, M. (2004). Outdoor ecology education and pupils' environmental perception in preservation and utilization. *Science Education International, 15*, 27-48.

Bozkurt Altan, E., Üçüncüoğlu, İ., & Öztürk N. (2019). Preparation of out-of-school learning environment based on STEM education and investigating its effects. *Science Education International, 30*(2), 138-148.

Büyüköztürk, Ş., Çokluk, Ö., & Köklü, N. (2011). *Sosyal bilimler için istatistik* (7. baskı). Pegem Akademi.

Bybee, R. W. (2009). *The BSCS 5E instructional model and 21st century skills.* Colorado Springs.

Cooper, R., & Heaverlo, C. (2013). Problem solving and creativity and design: What influence do they have on girls' interest in STEM subject areas?. *American Journal of Engineering Education, 4*(1), 27-38.

Chin, C-C. (2004). Museum experience- A resource for science teacher education. International *Journal of Science and Mathematics Education, 2*, 63-90.

Cunningham, C. M., & Kelly, G. J. (2017). Epistemic practices of engineering for education. *Science Education, 101*(3), 486–505. https://doi.org/10.1002/sce.21271

Çelik, İ. (2012). *Bir bilim kampından notlar. TUBİTAK Bilim ve Teknik Dergisi,* 538, 15-19.

Çevik, M. (2017). Ortaöğretim öğretmenlerine yönelik FeTeMM farkındalık ölçeği (FFÖ) geliştirme çalışması. *Journal of Human Sciences, 14*(3), 2436-2452.

Çevik, M., & Abdioğlu, C. (2018). Bir bilim kampının 8. sınıf öğrencilerinin STEM başarılarına, fen motivasyonlarına ve üstbilişsel farkındalıklarına etkisinin incelenmesi. *İnsan ve Toplum Bilimleri Araştırmaları Dergisi, 5*(7), 304-327.

Çevik, M., & Senturk C. (2019). Multidimensional 21th century skills scale: Validity and reliability study. *Cypriot Journal of Educational Sciences, 14*(1), 011–028.

Çorlu, M. S., Capraro, R. M., & Capraro, M. M. (2014). Introducing STEM education: implications for educating our teachers for the age of innovation. *Eğitim ve Bilim, 39*(171), 74-85.

Grajcevci, A., & Shala, A. (2016). Formal and non-formal education in the new era. *Action Researcher in Education, 7*(7), 119-130.

Greenwood J. (2020) On two foundational principles of the Berlin school of Gestalt psychology. *Review of General Psychology, 24*(3), 284–294. https://doi.org/10.1177/1089268019893972

Guba, E. G., & Lincoln, Y. S. (1982). Epistemological and methodological bases of naturalistic inquiry. *Educational Communication and Technology Journal, 30*(4), 233-252.

Hamarat, E. (2019). *In focus on 21st century skills Turkiye's education policy.* SETA Analiz, No. 272.

Holloway, I., & Wheeler, S. (1996). *Qualitative research for nurses.* Blackwell Science Ltd.

Kalemkuş, J. (2021). Fen bilimleri dersi öğretim programı kazanımlarının 21. yüzyıl becerileri açısından incelenmesi. *Anadolu Journal of Educational Sciences International, 11*(1), 63-87. https://doi.org/10.18039/ajesi.800552

Karademir, E. (2018). Okul dışı ortamlarda fen öğretimi. O. Karamustafaoğlu, Ö. Tezel ve U. Sarı (Eds.), *Güncel yaklaşım ve yöntemlerle etkinlik destekli fen öğretimi* (p. 426- 447). Pegem Akademi.

Karisan, D., Macalalag, A., & Johnson, J. (2019). The effect of methods course on preservice teachers' awareness and intentions of teaching science, technology, engineering, and mathematics (STEM) subject. *International Journal of Research in Education and Science, 5*(1), 22-35.

Koosimile, A. T. (2004). Out-of-school experiences in science classes: problems, issues and challenges in Botswana. *International Journal of Science Education, 26*(4), 483 – 496.

Krefting, L. (1991). Rigor in qualitative research: The assessment of trustworthiness. *The American Journal of Occupational Therapy, 45*(3), 214-222.

Kubat, U. (2018). Opinions of pre-service science teachers about outdoor education. *Mehmet Akif Ersoy University Journal of Education Faculty, 48,* 111-135.

Laçin Şimşek, C. (Ed.). (2011). *Fen öğretiminde okul dışı öğrenme ortamları*. Pegem Akademi Yayınları.

Larson, L. C., & Miller, T. N. (2011). 21st century skills: Prepare students for the future. *Kappa Delta Pi Record, 47*(3), 121-123.

Markowitz, D. G. (2004). Evaluation of the long-term impact of a university high school summer science program on students' interest and perceived abilities in science. *Journal of Science Education and Technology, 13*(3), 395-407.

Martin, L. M. W. (2004). An emerging research framework for studying informal learning and schools. *Science Education, 88*(S1), 71-82.

Miles, M, B., & Huberman, A. M. (1994). *Qualitative data analysis: An expanded Sourcebook. (2nd ed).* Sage.

Morse, J. M. (1991). Approaches to qualitative-quantitative methodological triangulation. *Nursing Research, 40*(1), 120–123.

National Science Teachers Association [NSTA] (2011). *Quality Science Education and 21st Century Skills*. Retrieved from http://science.nsta.org/nstaexpress/PositionStatementDraft_21stCenturySkills.pdf

Ngaka, W., Openjuru, G., & Mazur, R. E. (2012). Exploring formal and non-formal education practices for integrated and diverse learning environments in Uganda. *The International Journal of Diversity in Organizations, Communities and Nations, 11*(6), 109- 121.

OECD. (2019). *Educating 21st century children: Emotional well-being in the digital age (t. burns ve f. gottschalk, ed.). organisation for economic co-operation and development.* https://doi.org/10.1787/b7f33425-en

Orion, N., Hofstein, A.,Tamir, P. & Giddings, G. J. (1997) Devolopment and validation of an instrument for assesing the learning environment of outdoor science activities. *Science Education, 81*, 161-171.

Özbilen, A. G. (2018). STEM eğitimine yönelik öğretmen görüşleri ve farkındalıkları. *Bilimsel Eğitim Araştırmaları Dergisi, 2*(1), 1-21.

Özçelik, A., & Akgündüz, D. (2018). Evaluation of gifted/talented students' out-of-school STEM education. *Trakya University Journal of Education Faculty 8*(2), 334-351.

Panizzon, D., & Gordon, M. (2003). Mission possible: a day of science, fun and collaboration. *Australian Primary Junior Science Journal,19*(2), 9-14.

Peterson, T., & Fix, S. (Eds.). (2007). *Afterschool advantage: Powerful new learning opportunities*. Moorestown.

Plano Clark, V. L., & Creswell, J. W. (2008). *The mixed methods reader*. Sage.

Privitera, G. J. (2015). *Statistics for the behavioral sciences* (2nd edition). Sage Publications.

Randler, C., Kummer, B., & Wilhelm, C. (2012). Adolescent learning in the zoo: Embedding a non-formal learning environment to teach formal aspects of vertebrate biology. *Journal of Science Education and Technology, 21*, 384-391.

Schnittka, C. G., Bell, R. L., & Richards, L. G. (2010). Save the penguins: Teaching the science of heat transfer through engineering design. *Science Scope, 34*(3), 82–91.

Scott Simmons, O. (2021). A holistic model for student success in STEM (with J. Adams, D. Bright, J. Jackson), in social justice and education in the 21st century: Research from South Africa and The United States (Willie Pearson & Vijay Reddy eds.) (Springer 2021). Available at SSRN: https://ssrn.com/abstract=3829944

Sharts-Hopko, N. C. (2002). Assessing rigor in qualitative research. *Journal of the Association of Nurses in Aids Care, 13*(4), 84-86.

Shenton, A. K. (2004). Strategies for ensuring trustworthiness in qualitative research projects. *Education for information, 22*(2), 63-75.

Smith, J. W. (2017). Outdoor education. In *The Good Education of Youth* (pp. 366-367). University of Pennsylvania Press.

Sullivan, F. R. (2008). Robotics and science literacy: Thinking skills, science process skills, and systems understanding. *Journal of Research in Science Teaching, 45*(3), 373-394.

Sungur Gül, K., & Marulcu, İ. (2014). Investigation of in service and preservice science teachers' perspectives about engineering-design as an instructional method and legos as an instructional material. *International Periodical for The Languages, Literature and History of Turkish or Turkic, 9*(2),761-786.

Tabachnick, B., & Fidell, L. (2007) *Using multivariate statistics.* Boston: Allyn & Bacon.

Tatar, N. & Bağrıyanık, K. E. (2012). Opinions of science and technology teachers about outdoor education. *Elementary Online, 11*(4), 882-896.

The Parlimentary Office of Science Techonogy. (2011). *Informal STEM Education.* Retrieved from https://post.parliament.uk/research-briefings/post-pn-382/

Thomas, G. (2010) Facilitator, teacher, or leader? Managing conflicting roles in outdoor education. *Journal of Experiential Education, 32*(3), 239–254.

Timur, S., Timur, B., Yalçınkaya-Önder, E., & Küçük, D. (2020). Attitudes of the students attending out-of-school STEM workshops towards STEM education. *Journal of Theoretical Educational Science, 13*(2), 334-351.

Vygotsky, L. S. (1962). Thought and language. Cambridge, MA: MIT Press.

Wang, Hui-Hui. (2012). A new era of science education: science teachers' perceptions and classroom practices of science, technology, engineering and mathematics (STEM) integration. Retrieved from the University of Minnesota Digital Conservancy, https://hdl.handle.net/11299/120980

Wasti, S. P., Simkhada, P., Van Teijlingen, E. R., Sathian, B., & Banerjee, I. (2022). The Growing Importance of Mixed-Methods Research in Health. *Nepal Journal of Epidemiology*, *12*(1), 1175-1178.

Wilcox, R. R. (2012). *Modern statistics for the social and behavioral sciences: A practical introduction.* Chapman & Hall/CRC Press.

Yıldırım, A., & Şimşek, H. (2013). *Sosyal bilimlerde nitel araştırma yöntemleri (9. Baskı).* Seçkin Press.

Young, J., Ortiz, N., & Young, J. (2017). STEMulating interest: A meta-analysis of the effects of out-of-school time on student STEM interest. *International Journal of Education in Mathematics, Science and Technology*, *5*(1), 62-74.

Journal of Theoretical Educational Science, 17(1), 80-99, January 2024
Kuramsal Eğitimbilim Dergisi, 17(1), 80-99, Ocak 2024
[Online]: http://dergipark.org.tr/akukeg
DOI number: http://doi.org/10.30831/akukeg.1331720

Collegial Practicum Journey of EFL Teacher Trainees through Vignette-based Reflections

İngilizce Öğretmen Adaylarının Vinyet Tabanlı Yansıtmalar Aracılığı ile Meslektaş Dayanışmalı Staj Yolculuğu

Burcu TURHAN*

Received: 23 July 2023 **Research Article** **Accepted:** 23 October 2023

ABSTRACT: With the intent of facilitating the professional development of English-as-a-foreign language (EFL) teacher trainees, the impetus of this qualitative study is to evaluate the depth and breadth of teacher reflection using vignettes within an initial teacher education program. To this end, this study utilized the vignette technique to reveal 12 EFL teacher trainees' reflectivity from a collegial perspective. Over a nine-week time period, the participants worked in pairs, and each pair first created vignettes based on their observations or teachings in practicum. Afterwards, they generated reflective responses to the pre-designed questions, which led them to think about the problematic or critical issues raised in the vignettes of their pairs. At the end of the process, a total of 108 vignettes were gathered. The whole qualitative data was exposed to the analysis in light of the two-dimensional framework of Fund et al. (2002), which focuses on both the depth and breadth of teacher reflection. Results indicate that the vignette-based reflections showed more balanced variety in breadth, whereas they did not contain a great number of comments made at deeper levels of reflection. Therefore, it is recommended that all the stakeholders should take more active roles to foster reflectivity and criticality toward educational matters raised in the vignettes.

Keywords: Reflection, vignette, teacher trainees, practicum, teacher education.

ÖZ: İngilizce öğretmen adaylarının mesleki gelişimine katkı sağlamak amacıyla tasarlanan bu nitel çalışmada bir öğretmen eğitimi programı kapsamında öğretmen adaylarının yansıtıcı fikirlerinin derinliği ve genişliği vinyet tekniği ile değerlendirilmiştir. Bu amaçla 12 İngilizce öğretmen adayının meslektaş dayanışması bağlamında yansıtıcılığını ortaya çıkarmak için vinyet tekniği kullanılmıştır. Dokuz hafta boyunca katılımcılar çiftler halinde çalışmış olup her bir çift ilk olarak stajdaki gözlemleri veya öğretimleri ile alakalı vinyetler oluşturmuştur. Daha sonra, her bir katılımcı çift olarak çalıştığı katılımcının oluşturmuş olduğu vinyetlerde ortaya çıkan sorunlu veya kritik konular üzerinde düşünmelerini sağlayan önceden belirlenmiş sorulara yansıtıcı yanıtlar üretmişlerdir. Bu süreç sonunda toplam 108 adet vinyet elde edilmiştir. Tüm nitel veriler, Fund ve diğerlerinin (2002) öğretmen yansıtmasının hem derinliğine hem de genişliğine odaklanan iki boyutlu çerçevesi kullanılarak analiz edilmiştir. Sonuçlar, vinyet tabanlı yansıtmaların genişlik boyutunda daha dengeli bir çeşitlilik gösterdiğini, buna karşın daha derin yansıtma düzeylerinde yapılan çok sayıda yorumu içermediğini göstermektedir. Bu nedenle, tüm paydaşların vinyetlerde dile getirilen eğitim konularına yönelik yansıtma ve eleştirelliği teşvik etmek adına daha aktif roller almaları önerilmiştir.

Anahtar kelimeler: Yansıtma, vinyet tekniği, öğretmen adayları, staj, öğretmen eğitimi.

* Asst. Prof. Dr., Hatay Mustafa Kemal University, Hatay, Türkiye, burcu.oyp@gmail.com, https://orcid.org/0000-0001-6025-0616

Citation Information
Turhan, B. (2024). Collegial practicum journey of EFL teacher trainees through vignette-based reflections. *Kuramsal Eğitimbilim Dergisi [Journal of Theoretical Educational Science], 17*(1), 80-99.

ISSN: 1308-1659

The theory and practice dilemma in teacher education has been discussed for a couple of decades. In order to eliminate the negative consequences of such a dilemma, teachers can be engaged in reflective practices in which they look at educational issues through critical lenses. It is vital for teacher training programs to offer reflective practices so that the gap between theory and practice can be minimized (Ibrahim-Didi, 2015). Indeed, not only the gap between theory and practice but also the complexity of teaching requires teachers to question their practices with the ultimate intent of promoting professional development as well as student performance. Thus, reflective practices should be at the forefront since it is only possible to benefit from the process of continuous learning with reflection. Therefore, it is highly crucial to incorporate reflective practices into both pre-service and in-service teacher training programs. The key rationale behind the integration of reflective practices throughout a teacher's career is that experience alone may not necessarily lead to learning; purposeful reflection on specific experience is a must (Mathew et al., 2017). A reflective teacher is the one who not only looks back on previous actions and events but also examines emotions, experiences, actions, and responses from a reflective stance. Relatedly, within the scope of this paper, reflection is regarded as a flashback that teachers need to mediate for their improvement (Mathew et al., 2017).

Reflective practice is defined as an activity that involves continuous examination of one's own experiences in practicing theory under the guidance of professionals (Schon, 1983). In teacher education contexts, those professionals are teacher educators/supervisors and mentors. With the guidance of these professionals, teachers and teacher trainees have the chance to analyze their feelings, evaluations, practices, experiences, and so forth (Gibbs, 1988). Further, reflective practice is linked to lifelong learning, which leads to the development of autonomous, qualified, and self-directed teachers who have a good level of expertise in the teaching profession (Jasper, 2003). More importantly, reflective teachers should be able to move beyond the basic concern of how-to questions about instructional methodologies; instead, they should ask what and why questions about instructions and managerial techniques as a part of a broader educational end (Bartlett, 1990). Asking what/why questions is essential because such awareness helps teachers manage their teaching effectively, resulting in the emergence of autonomous teachers. From Dewey's (1933) point of view, reflection is a cyclical and iterative process in which one deals with solving problems. In this process, teachers become more aware of their experiences and apply professional knowledge to these experiences (Zeichner & Liston, 2013). Osterman and Kottkamp (2015) highlight the dialogic nature of reflective practice in which teachers can move back and forth among past, present, and future events by generating various perspectives on the same practices or observations. To add more, this dialogic travel can be accepted as a form of self-evaluation which flourishes the production of new opinions and learning/teaching modes (Boud, 2001; Moon, 2004). Apart from these reflective self-evaluations, peer sharing and observations are mostly preferred by teachers owing to their time-saving nature in a busy school schedule (Kharlay et al., 2022). Particularly for foreign language teaching, EFL teacher trainees are found to believe in the power of reflection for their professional development (Turhan & Kırkgöz, 2018), and there are certain slight changes in their views toward language teaching as a result of reflection (Turhan & Kırkgöz, 2023). Yet, EFL teacher trainees could not reflect at a higher level of

criticality in time as they took part in systematic and regular reflective practices (Turhan & Kırkgöz, 2018, 2023).

Being a significant element of initial teacher training programs, reflective practice could be utilized as a way to renew teacher trainees' teachings and realizations of the impacts of their teaching (Jacobs et al., 2011), especially when they evaluate whether their teaching is sound and opens doors for fruitful learning. Akbari (2007) supports this statement by suggesting that reflective practice allows teachers to question clichés learned in the first years of their career; hence, they become able to design and implement more informed teaching practices. Especially for teacher trainees, it is important to question the clichés so that they can later recognize multiple possibilities that they can employ in their future classrooms. However, to accomplish this kind of questioning, teachers need to have critical reflection skills that do not develop naturally or by automatic occurrence for many of them (Yang, 2009). Moreover, despite all the aforementioned positive ties between teaching and reflection, it is a fact that not every reflective practice guarantees effective teaching (Boud & Walker, 2015). If the objective is to create changes in teachers' actions or thinking, we should question the quality of reflection in order to make sure that reflection leads to adequate judgment in teachers' actions (Yılmaz & Akar, 2022). Likewise, Davis (2006) distinguishes between productive and unproductive reflection. She considers reflection to be unproductive if it fails to analyze teaching by backing claims with evidence, questioning assumptions, and considering alternatives. In fact, evolving into a qualified reflective teacher is similar to a skill development process. For this reason, learning reflective skills necessitates a classroom designed as a kind of laboratory where teacher trainees can relate theory to practice in the accompaniment of teacher educators. In those laboratories, the potential reflective techniques are typically reflective diaries, collaborative learning, recording lessons, peer observation, educators' feedback, action research, reverse mentoring, and so on. Unfortunately, the appropriateness and effectiveness of such instruments in facilitating quality reflection is unclear (Ambler, 2012; Stecher et al., 2006). As an alternative to those typical techniques, as in the current study, teacher educators can also adapt the use of vignettes as a reflective tool during practicum. Using vignette-based instruments is a more suitable way of obtaining important clues on the quality of reflection (Jeffries & Maeder, 2011), and this notion shapes the main logic behind the integration, examination, and assessment of vignettes in this study.

Vignettes are short, authentic stories that can be employed in differing formats, such as written documents, videos, or even cartoons, all of which trigger thinking and discussion (Henderson et al., 2016). Jeffries and Maeder (2011) define vignettes as a sort of short story describing an identified problem. They additionally argue that vignettes are useful because they are brief and easy to construct and administer, present a well-designed stimulus for critical thinking, are valuable in dealing with sensitive topics, and can be used with individuals and groups in both face-to-face and online settings. They also list the four criteria for defining vignettes: (1) the story describes a problem in a maximum of 200 words; (2) it simplifies a real-life/real event in a way that any other individual can understand even though they do not have expertise in the field; (3) its set of tasks/questions is connected to a scoring scheme; and (4) it is deliberately incomplete so that multiple solutions to the story's problem can be offered. As it is clear, vignettes are, if used in the right way, reliable tools for rating what is learned.

They can be utilized as a method for creating open-ended discussion sessions where participants explore topics and share multiple views. Topics of vignettes might be diverse, ranging from professional practices with a problem to excellence in teaching (Jeffries & Maeder, 2011). Examining other definitions of vignettes in the literature, a vignette is an effective written or visual stimulus to which participants are invited to answer (Hughes & Huby, 2004). Vignettes are incomplete short stories used for reflecting on real-life situations, and they gauge discussions in order to produce multiple possible solutions to problems (Jeffries & Maeder, 2005). In Torres's (2009) words, vignettes depict specific cases involving problems with the intent of probing individuals about the way they understand these. In other words, vignettes primarily illustrate descriptive episodes associated with real incidents and are presented either in a written or visual format (Skilling & Stylianides, 2020). Even though visual vignettes can generate rich information (Bradbury-Jones et al., 2014) and produce more trustworthy data (Torres, 2009), the written formats of vignettes can be advantageous since they might include dialogic texts, comic annotations, or mathematical representations (Friesen & Kuntze, 2016, 2018) at the same time.

Vignettes have been used for various research purposes (e.g., Bradbury-Jones et al., 2014; McGarr & Gallchóir, 2020; Skilling & Stylianides, 2020; Tasar, 2006; Torres, 2009; Volkmann, 2000). To exemplify, they are found to protect research participants by offering chances to take the role of another person (Bradbury-Jones et al., 2014), and thus, they can provide more reliable data about sensitive issues. Similarly, asking teacher trainees about educational policies might prevent them from sincerely putting their views into words. Considering this, McGarr and Gallchóir (2020) create a realistic context through the use of vignettes, thanks to which they can elicit the teacher trainees' justifications about the policies of technology integration into education. Beyond these, Torres (2009) claims that vignettes help one interpret value-laden conceptions. As one of the value-laden conceptions, teacher beliefs can be investigated through the use of purposefully constructed vignettes (Skilling & Stylianides, 2020). Tasar (2006) claims that teacher trainees' scientific knowledge can be evaluated with carefully chosen vignettes. Besides, Volkmann (2000) suggests that vignettes are helpful tools to create an environment for teacher trainees to apply educational theory to their own teaching experiences and observations. Similar to the procedure followed for the EFL teacher trainees in this study, Wilkerson et al. (2018) apply a Vignette Activity Sequence in which teacher trainees are first acquainted with a vignette, and second, they write vignettes, unravel that it is a beneficial procedure to help trainees make critics on effective teaching practices in the field of math. Some other studies employ vignettes with in-service teachers and conclude that teachers could also benefit from vignettes for professional development (Ambler, 2012; Angelides & Gibbs, 2006; Jeffries & Maeder, 2005). Depending on the above summary, the value of the present study is to encourage teacher trainees to design their own vignettes and make the analyses collaboratively rather than analyzing ready-made vignettes individually. Such a procedure makes vignettes revolve around the real educational issues encountered and detected in practicum. In parallel with this, the following research questions guide the study:

a) What are the depth and breadth of vignette-based reflections written by the EFL teacher trainees?

b) Does the depth and breadth of the EFL teacher trainees' vignette-based reflectivity develop as they progress throughout the practicum? If yes, how?

Method

From a qualitative standpoint, this study was conducted using the vignette technique, which relates research to practice and results in transformative experiences by practitioner-researchers (Nind & Pepin, 2009). The main rationale for the use of the vignette technique is to better understand the participants' interpretation of specific situations in the light of their subjective views, especially in educational settings where the interconnections among emotional, social, and cultural factors are strong (Stecher et al., 2006). A written format of vignettes was preferred in the current study because they provide more variety and details in relation to observed or experienced issues. For the sake of ethical considerations, the necessary permission to conduct the study was obtained from Hatay Mustafa Kemal University Social and Human Sciences Scientific Research and Publication Ethics Committee with the 902-01-FR 006 document number (07.04.2023).

Research Context and Participants

The research site of this study is a state university in the southern part of Turkey. Within the faculty of education, the students studying in the English Language Teaching (ELT) Program are under scrutiny. This program offers teacher trainees a four-year undergraduate study. After students are enrolled in the program, depending on their scores in the University Entrance Examination, they receive a language proficiency exam, and if they successfully pass this exam, they start the ELT department as freshmen. If they fail the exam, they have to take part in the preparation class for one academic year, and they receive a language proficiency exam at the end of the preparation class so that they can start their first year of study in the department. The current teacher education curriculum implemented in the department includes both theoretical and practical courses; in the first three years, theoretical courses such as ELT Approaches, Language Acquisition, Teaching Language Skills, Teaching English to Young Learners and English Literature are dominant. In the last year of study, the Teaching Practice course gives teacher trainees chances to make teaching practices in real English classes at practicum schools for two semesters. The current study was carried out in the scope of this Teaching Practice course, and a total of 12 EFL teacher trainees (four females and eight males) attended the practicum schools, which were determined by the Ministry of National Education during the spring semester of the 2022-2023 academic years. They attended either a secondary or a high school for six class hours (240 minutes) a week. Both of the practicum schools were located in the city center and can be regarded as advantageous schools in terms of educational facilities at the time of the study. The participants were between the ages of 22-23 and had neither previous teaching experiences, apart from the micro-teaching sessions in the third year of study at university, nor reflective practice experiences.

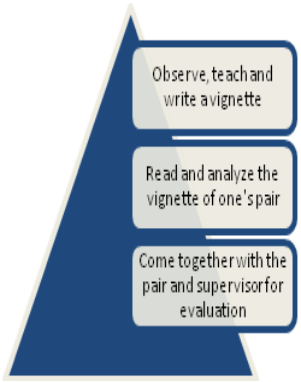
Procedures

The research procedure was organized in parallel with the steps followed in the Teaching Practice course on a weekly basis. However, the way of teaching was

organized and applied differently from its traditional way of teaching, which includes teacher trainees' individual work evaluated weekly by mentors and supervisors in light of an official observation checklist. Instead, during this study, 12 EFL teacher trainees worked in pairs; namely, six pairs were always in an attempt to provide collegial support to one another during practicum. In the first week, the participants received an introductory session to the Teaching Practice course and its core requirements. In the second and third weeks, they were informed about the content and the method that would be utilized during the course. More specifically, they collaborated with the supervisor with the intent of developing a vignette-based instrument (see details in the subsection of data collection instrument**)** to be employed for reflective purposes. In the remaining nine weeks, they followed the steps illustrated in Figure 1.

Figure 1

The Steps Taken During the Teaching Practice Course for Nine Weeks



As evident in Figure 1, the first task of the participants was to carry out observations and practice teaching at practicum schools by paying attention to their mentors' comments and advice. Afterwards, they were required to write a vignette that presented a critical or problematic issue encountered during the practicum day each week. Their second task was to share the vignettes with the pairs so that the pairs could analyze the issue/s raised in the vignette through critical and reflective lenses. Lastly, their third task was to be involved in an evaluation session with the pairs and the supervisor. In those evaluation sessions, the objective was to look back on what happened that week and share ideas about the knowledge gained with the help of the vignette writing process, followed by vignette-based reflections. As understood from Figure 1, the procedure was a cumulative knowledge-shaping process, and each week's learning outcomes were an opportunity to gain a broader and deeper understanding of the upcoming week's outcomes.

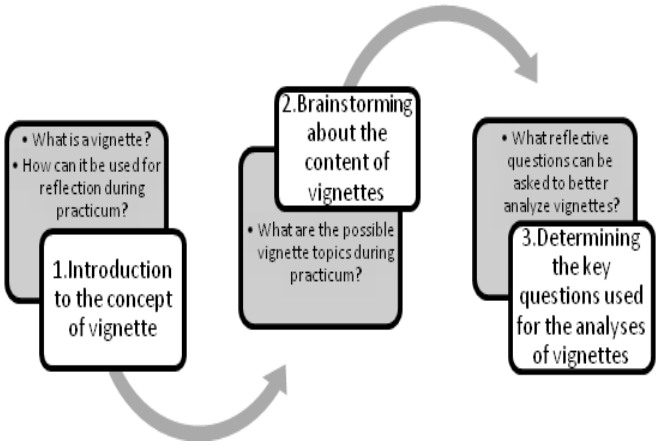
Data Collection Instrument

The vignette-based instruments are normally generated or adapted by the teacher educators or instructors (Jeffries & Maeder, 2011) and are not prepared by teacher trainees or students themselves. Yet, in this study, teacher trainees were asked to prepare their own vignettes based on their observations and experiences during practicum. This could be accepted as an innovative approach in that sense since EFL teacher trainees took two roles as both vignette generators and vignette analyzers. As clarified in the previous subsection, after the introductory week, the instrument

generation was finalized in two weeks. The flow of the instrument generation process is demonstrated in Figure 2.

Figure 2

The Generation Process of the Vignette-based Instrument



Based on Figure 2, the details concerning the collaborative generation process of the vignette-based instrument are as follows:

Week 1: At the beginning of the term, the teacher trainees became familiar with the different definitions and samples of vignettes from different teacher education fields. Additionally, how vignettes are used as a professional teacher development activity was discussed. It was also decided that vignettes would be mainly written, but it was possible for the participants to add quotes, cartoons, or images into their vignettes so as to enrich the meaning they wanted to convey.

Week 2: In a brainstorming session, the teacher trainees and their supervisor, who was also the researcher, reached a consensus on the possible observation topics, some of which were teaching context, school facilities, the existing curriculum, dominant teaching methods, student characteristics, subject matter, organization of a typical English lesson, lesson delivery, materials, building rapport, classroom management, feedback, assessment, physical aspects of the classroom, the role of teacher trainees, teacher identity, and expectations from teacher trainees.

Week 3: The supervisor shared information about the varying reflection levels in terms of depth and breadth, the nature of reflective questions, and how to produce reflective responses to the issues or problems raised in the vignettes. To achieve this, the supervisor presented sample responses to reflective questions prepared for an example vignette for the participants. Following this, the participants prepared these questions for their own vignettes: (a) What are the issues or problems raised in the vignette? Why do you think so?, (b) What would you do in a similar case as a teacher? How would you cope with the issues or problems?, (c) How do you feel about the issues or problems raised in this vignette? Besides, the supervisor provided more focused reflective questions during the process: (a) Would you personally be in favor of your pair's way of dealing with the problem?, (b) Could you suggest a different rationale for your

argument?, (c) Should teachers take a similar stance with you?, (d) Did you experience or observe this situation in your own teaching as well?, (e) Could you find any ties between this situation and what you have learnt in the methodology courses at university?

Data Analysis Framework

In evaluating teacher reflection, the usual way is to resort to a hierarchical typology that starts from a low level of reflection (e.g., making technical and descriptive comments) and moves towards a higher level (e.g., considering moral and ethical rationales). To exemplify, Jay and Johnson (2002) regard reflection as a decision-making process composed of three core stages, which are description, comparison, and criticism. Later, Larrivee (2008) suggests the categories of pre-reflection, surface reflection, pedagogical reflection, and critical reflection for scrutinizing teacher reflection. In addition, Kayapinar and Erkus (2009) refer to three dimensions of reflection, ranging from being non-reflective and reflective to being critically reflective. The danger in such typologies is reducing teacher reflection to one dimension, considering a set of rigid and pre-packaged categories or stages. To eliminate such a danger, in this study, a two-dimensional framework designed to evaluate written reflective tasks in teacher training courses (Fund et al., 2002) was utilized. This framework's first dimension is related to the breadth of reflection, and the second dimension is for exploring the depth of reflection. For the breath of reflection, there are three cells (subject-matter content, didactic content, and personal content) in it. *Subject-matter content* is linked to what question and the educational issues debated in the vignette. *Didactic content* deals with how-to questions, strategies and methods to teach or learn the content. *Personal content* reveals one's views toward themselves (the I) as a teacher. Moreover, for the depth of reflection, there are four cells (description, personal opinion, linking, and critical bridging) in the framework. *Description* level requires one to describe the event or case without any further comment or evaluation. *Personal opinion* level is associated with one's own non-theoretical stance rather than attaching importance to evidence from the relevant literature. *Linking* concerns the bonds between educational issues and previous knowledge or literature without any further elaboration. *Critical bridging* necessitates one to deliberately discuss and find possible alternative opinions from the literature, as well as to decide how to take action in the future. In light of this framework, the analyses were conducted twice by the researcher so as to ensure intra-rater reliability and member check was achieved through telephone calls. For each cell in the framework, sample evidence from the collected data is outlined in the Appendix.

Results

The Depth and Breadth of EFL Teacher Trainees' Vignette-Based Reflections

Vignette-based reflections were scrutinized in a way that the reflectivity of the EFL teacher trainees throughout the practicum was explored with regard to the depth and breadth dimensions, as illustrated in the following graphs.

Figure 3

The Overall Depth of Reflections

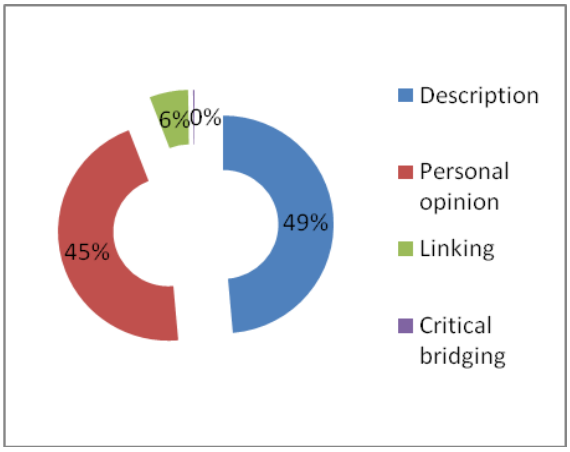
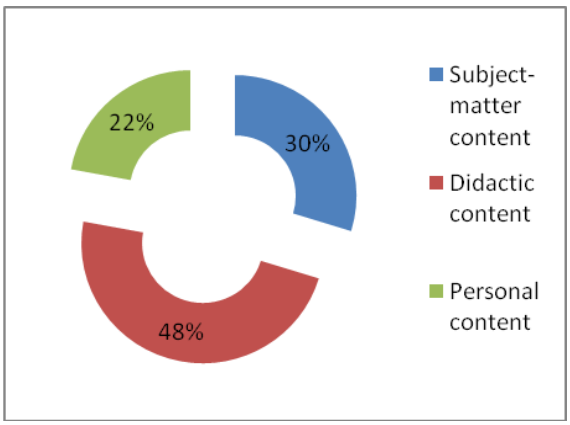


Figure 4

The Overall Breadth of Reflections



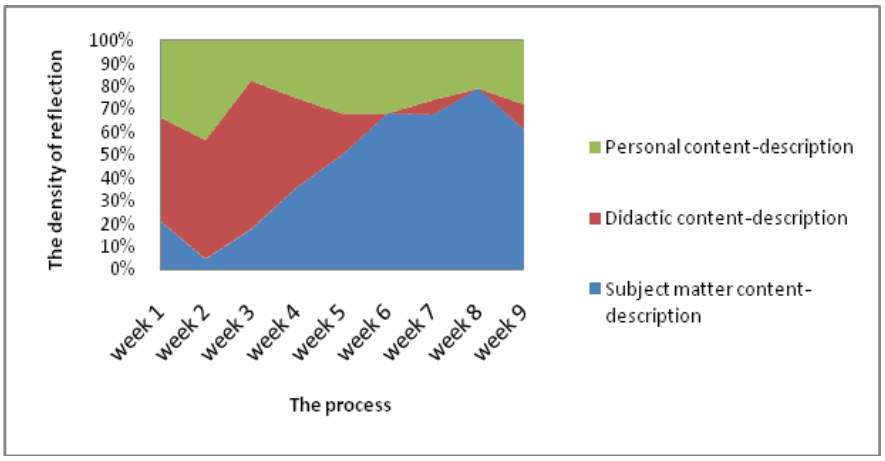
As it is clear in Figure 3, in terms of the depth dimension, the participants tended to write reflections on both their own vignettes and responses to their pairs' vignettes at the level of either description or personal opinion most of the time. This shows that their reflection tendency mostly led them to describe or share personal concerns about what was happening in the teaching/learning environment. As for the other levels, it is obvious that the participants almost never resorted to critical bridging in their reflections, whereas they reflected at the level of linking a few times. That is, it is unlikely to assert that the participants were able to exhibit deep reflective thinking skills. This is an indication of their unawareness or inability to reflect on educational issues encountered in practicum through a comprehensive analysis of reasons, rationales, and alternatives or building associative ties to previous scientific knowledge. Further, regarding the breadth dimension, Figure 4 shows that the employment of different reflection levels indicated a more balanced distribution contrary to the depth dimension. In other words, the participants wrote about didactic issues the most, yet they paid attention to subject matter and personal issues to a certain extent as well. To compare, it is evident that the participants' focus on themselves as a human, a student, or a teacher did not seem to be as dominant as their focus on what and how lessons are taught in their reflections.

The Reflectivity Flow of EFL Teacher Trainees throughout the Practicum

In order to unravel whether the vignette-based reflectivity of the participants develops as they practiced teaching during the practicum, the obtained data were examined on a weekly basis, considering the depth and breadth dimensions of reflection. The results are presented for each reflection level included in the depth dimension, one of which is demonstrated in Figure 5.

Figure 5

The Reflectivity of the Participants for the Description Level on a Weekly Basis



Based on Figure 5, the participants were inclined to describe the issues in relation to subject-matter content more frequently, beginning from week 3 to the end of the process. In the first five weeks, they generally reflected on the didactic content; however, reflections on the subject matter and personal content seem to be much more dominant in their vignettes toward the end of the process. Interestingly, the density of descriptive reflection on the subject matter and didactic content fluctuates in opposite directions. Namely, the references to didactic content decrease as the references to subject-matter content increase or vice versa. Yet, the density of personal content appears to be more balanced in the reflections throughout the nine-week period, even though the participants did not draw attention to personal issues as much as they did for the didactic and subject-matter issues on the whole. The below excerpts exemplify the descriptions of subject matter, didactic, and personal content consecutively.

> "With 6th graders, I conducted an activity related to the unit of At the Fair. [The trainee explains that they listen to a song and complete the missing words in the lyrics of the song.] Then, I continued with a vocabulary quiz to check whether students could comprehend concepts about being in a fair." (Week 9, P5)

> "The teacher struggled for 10 minutes to ensure silence in the classroom. For some reason, all the students were very active today. Later, the teacher listened to the complaints and desires of the children who raised their hands." (Week 3, P4)

> "It is very sad to observe that students are reluctant and do not bring the necessary materials to the class. Unfortunately, such setbacks demotivate me as a future teacher." (Week 7, P2)

In accordance with the level of personal opinion, Figure 6 represents to what extent the participants reflected on the subject matter, didactic, and personal content in light of their current intuitions, feelings, or insights during the whole process.

Figure 6

The Reflectivity of the Participants for the Personal Opinion Level on a Weekly Basis

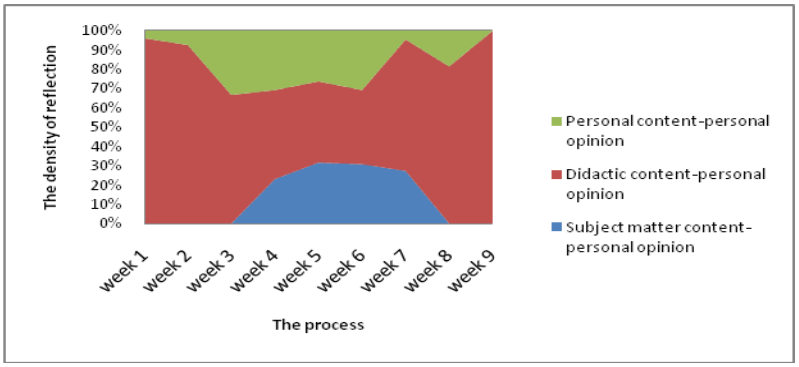


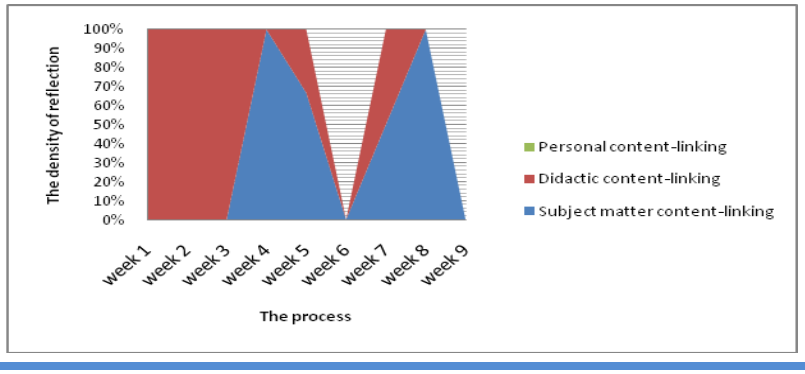
Figure 6 clearly shows that the participants' reflections at the level of personal opinion are densely full of their references to didactic content that they observed or taught in practicum. In fact, the density of didactic content seems to be less between weeks 3 and 8. This is most probably because of their tendency to write about subject matter or personal issues during this time period. It should also be noted that the participants did not share their personal views or comments about subject-matter content at the beginning and at the end of the process. Besides, they reflected on personal issues at the level of personal opinion more or less throughout all weeks, even if this type of reflectivity did not dominate all of their reflections. The related examples of personal opinions on subject matter, didactic, and personal content are presented respectively in the excerpts below.

> "This week, the teacher used an extra activity related to the unit of Hobbies. It seemed really enjoyable at first, but I realized that most students were unfamiliar with the content of the activity. This made things complicated for them." (Week 6, P8)
>
> "I think that children should be approached considering their ages. By doing so, we can reach better learning outcomes because children need interest, care, and concern." (Week 1, P12)
>
> "It was annoying to witness that students had a communication problem, in fact, a misunderstanding. At first, I had no idea about how to react, and for this reason I did want to be a part of this. Later, I thought I should intervene in the situation because students see us as a guide who leads them to the right road." (Week 6, P5)

Following this, a different type of fluctuation is revealed for the level of linking, as pictured in Figure 7.

Figure 7

The Reflectivity of the Participants for the Linking Level on a Weekly Basis



Depending on Figure 7, it is possible to state that the participants never resorted to the level of linking when they reflected on personal issues, which means they did not have an explicit awareness of the linking and connecting processes they were going through. However, they were involved in reflections which were concentrating on didactic and subject-matter issues. That is to say, they were capable of creating connections between what and how lessons are taught and prior knowledge acquired from the relevant literature or teacher training courses. It is also understood from Figure 7 that the participants could suggest possible reasons for didactic issues in light of previously learned concepts or theories more frequently than they could for subject-matter issues. The following excerpts present examples from the data for reflections on only subject matter and didactic content at the level of linking, respectively, since there is no reference to reflections on personal content at the level of linking.

> "Teaching unmotivated students may be both tedious and difficult for teachers. I think as teachers, we have lots of opportunities to make sure our students have positive attitudes toward English. We can achieve this through the integration of English cartoons into our lessons so that students can gain cultural awareness. This is like culture-focused teaching." (Week 4, P3)

> "I would try to choose activities that would not create excessive competition in the classroom but rather strengthen the unity and relations of the students. The reason behind this is all about the social nature of language learning, which is also highlighted by socio-constructivist teaching techniques." (Week 3, P10)

As for the last level of reflection for the depth dimension, Figure 8 summarizes the extent to which the participants reflected at the level of critical bridging.

Figure 8

The Reflectivity of the Participants for the Critical Bridging Level on a Weekly Basis

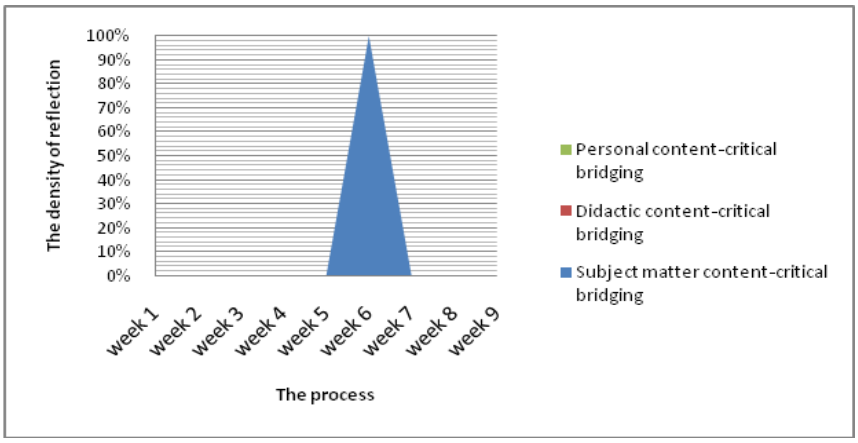


Figure 8 is an indication of the very rare employment of critical bridging by the participants in their vignette-based reflections. To be more precise, the participants reflected on subject-matter content at the level of critical bridging toward the end of the practicum process. Personal and didactic issues were not touched upon in this level of reflectivity even once. Overall, it should be emphasized that the participants were unaware of the fact that they could pose considerations, make judgments about the way of teaching, or critically analyze themselves in order to be more reflective teachers. Maybe they were just unable to do that. The only example of this level is evident in the below excerpt.

> "We have to use the mother tongue when teaching a foreign language to students at this age because cognitively immature students need to be exposed to content in their mother tongue [The teacher trainee suggests his own way of content design combining contents in the mother tongue and in English] Thus, students can connect previous and new concepts in their minds in a more structured and memorable way." (Week 6, P11)

Discussion and Conclusions

With the purpose of understanding the nature of EFL teacher trainees' reflections via the vignette technique, this study reveals how the breadth and depth of teacher reflection develop throughout a certain period of time. Based on the results, a great deal of the participants' reflections was found to be at the description and personal opinion levels despite the supervisor's weekly support of how to reflect at deeper levels. Namely, the analyses of the depth dimension showed that the participants did not reflect at deeper levels, which require one to generalize, give reasons, analyze from critical perspectives, and so forth. Identical findings in relation to the dominance of descriptive shallow comments in the EFL teacher trainees' reflections were reached in the studies of Turhan and Kırkgöz (2018, 2023). In terms of breadth, it is obvious that the participants could focus on multiple topics ranging from subject-matter content to didactic and personal content in their vignette-based reflections. Among these different content types, the least frequent references in the overall reflections were observed in personal content. This might be interesting and unexpected because the main aim of integrating a vignette-based reflection process into the practicum was to equip the participants with knowledge and awareness of oneself as a future teacher. Rather than this, the participants concentrated more on didactic content in their reflections, which pinpoints the fact that the way of teaching or learning drew much attention and needed to be analyzed more frequently. In fact, this could be a hopeful finding because reflection enhances informed teaching practices (Akbari, 2007) and is a road for renewing teaching as well as raising awareness toward the impacts of teaching (Jacobs et al., 2011). Additionally, the flow of reflectivity during practicum did not show a specific or meaningful change. Normally, it is expected that the participants become more reflective toward the end of the practicum process as they reflect more and more in time. However, in this study, no systematic flow toward a more reflective or critical stance was detected as a result of the vignette-based reflections. This is most probably because they were not able to approach issues in a similar level of reflectivity, and the content of issues may be a determinant for the reflective skills of the participants, who were really novice and new to the teaching profession. Another reason might be that they did not have any previous teaching and reflection experiences, which they could use as a reference point for comparing and contrasting their prior conceptions with new realizations. In other words, becoming more reflective may not be directly or solely related to the participants' involvement in reflection over longer time periods. As Yang (2009) claims, critical reflection skills do not seem to be acquired naturally. Maybe the best way is to deliver explicit instructions on how to become reflective practitioners. This arises from the fact that reflective skills are crucial to bridge the gap between theory and practice (Volkmann, 2000). Even so, it is valuable to be a part of such vignette-based reflection processes in the sense of assessing the scientific knowledge of teachers (Tasar, 2006) to ensure the quality of reflection (Jeffries & Maeder, 2011) and building critical connections to effective teaching practices (Wilkerson et al., 2018). The

reason behind incorporating any type of reflection is the desire to construct a quality base for teacher action and thinking (Yılmaz & Akar, 2022).

Suggestions

The main conclusions shed light on a few suggestions that could be beneficial for the supervisors or teacher educators who aim to design a similar collegial practicum process for teacher trainees, either in ELT or other teacher training programs. First, a great number of explicit examples for deeper levels of reflection were provided by the supervisor in the beginning weeks of this study, but the expected outcomes were not obtained in terms of criticality and reflectivity. Therefore, it would be better for the supervisors to be models of how to reflect more effectively with think-aloud protocols in which they exemplify each level of reflection concretely. By doing so, they can display how they direct and shape their thinking on diverse issues, especially for deeper levels of reflection. To achieve this, supervisors should ensure that teacher trainees have metacognitive thinking ability. Second, considering the breadth dimension, the results show that the least frequent content that teacher trainees were interested in was personal content, even though they mentioned personal issues at varying degrees throughout the practicum. To increase opportunities for looking back and forth on personal issues, supervisors could provide teacher trainees with guidance in relation to introspection more often. Third, in order to make such a process more collegial, mentors could be involved in vignette-based reflection processes since they are the models of teacher trainees in the field. For instance, mentors could also prepare vignettes covering issues observed around their own educational environments or lived teaching experiences in their own classes. Afterwards, they could share their vignettes with teacher trainees in regular face-to-face or online sessions organized for the critical analyses of vignettes in a collaborative way. Even, as a way to maximize reflection quality and create a sense of togetherness among colleagues, mentors can combine vignettes with such other self-development tools as digital narrative storytelling (e.g., Craig, 2013), video self-reflections (e.g., Arya & Christ, 2013; Calandra et al., 2014; Christ et al., 2012; Walshe & Driver, 2019) and video reflections with peers (Arya et al., 2015; Christ et al., 2012; Eröz-Tuga, 2013; Shanahan & Tochelli, 2014). All these tools have the potential to develop a greater understanding of trainees' strengths and weaknesses in a way that they expand their pedagogical knowledge through solidarity and collaborative professionalism. Thus, variety in the uses of reflective tools may trigger critical thinking more effectively because each trainee will most probably be successful at reflecting with different tools. Variety in this regard is also vital for boosting motivation and autonomy toward reflective practices by giving trainees chances to make their own choices among numerous reflection tools. More crucially, this could be an official procedure guided by supervisors, applied by mentors, and supported by the Ministry of National Education as an in-service professional development activity. This suggestion, if put into practice, can yield fruitful outcomes since some other studies also find vignette-based reflections beneficial for improving in-service teachers' expertise (e.g., Ambler, 2012; Angelides & Gibbs, 2006; Jeffries & Maeder, 2005). To add more, reflection is important for training qualified teachers, lifelong learning (Jasper, 2003), and professional and pedagogical development of teachers. In this regard, collegial support among teachers, trainees, supervisors, teacher educators, and mentors need to be encouraged in a well-designed vignette-based reflection practice. All in all, purposeful

reflection on specific experiences is essential (Mathew et al., 2017), and to accomplish this, vignette-based reflection may be incorporated into methodology courses offered in teacher training programs before the practicum experience. A further study could adopt a more collegial perspective and involve teacher trainees from different regions of Turkey or from different countries in a way that they collaborate for the analyses of their vignettes generated in different contexts through online platforms.

Limitations

This study is not without limitations. With its small-scale focus on vignette-based reflections of EFL teacher trainees, this study includes participants working in pairs, totaling 12 in number. To make the study more comprehensive, it would be sound to encourage all the senior students in the aforementioned ELT department to participate in the study so that the case of senior EFL teacher trainees at a state university could be investigated through the vignette technique utilized as a tool to explore content and criticality of reflection. With the use of vignettes, the participants searched for educational issues that were appropriate to be presented in a story-like format. However, this might cause them to miss or ignore other problematic issues that they believe were impossible or difficult to present in a story-like format. Hence, with the collaboration of trainees and supervisors, mentors might be informed about the details of the vignette-based reflection beforehand if they are not knowledgeable enough. This is because trainees could write better vignettes with the support of mentors. Last but not least, the sole data collection instrument used in this study was vignette-based reflections. Yet, it could be combined with other data collection instruments such as open-ended questionnaires or focus-group interviews. Thus, the views of all the parties about the process could be examined.

Acknowledgements

I would like to thank all the teacher trainees who provided data for this study.

Conflicts of Interest

There is no potential conflict of interest for this study.

Author Bio:

Burcu Turhan is an Asst. Prof. Dr. at English Language Teaching (ELT) Department of Hatay Mustafa Kemal University, Turkey. She holds her Master and PhD degrees from the ELT Department of Çukurova University. Her MA thesis is a study on English-medium instruction (EMI) at tertiary level, revealing motivational tendencies of engineering students and lecturers toward EMI. Besides, her PhD study focuses on fostering English language teacher trainees' immediacy behaviours via a specific training programme. Her general research interests include foreign language teacher education, pragmatics, EMI, metacognition, problem-based learning, discourse analysis and reflective practice.

References

Akbari, R. (2007). Reflection on reflection: A critical appraisal of reflective practice in L2 teacher education. *System*, *35*(2), 192-207. https://doi.org/10.1016/j.system.2006.12.008

Ambler, T. B. (2012). Autobiographical vignettes: A medium for teachers' professional learning through self-study and reflection. *Teacher Development, 16*(2), 181–197. https://doi.org/10.1080/13664530.2012.679864

Angelides, P., & Gibbs, P. (2006). Supporting the continued professional development of teachers through the use of vignettes. *Teacher Education Quarterly, 33*(4), 111–121.

Arya, P., & Christ, T. (2013). An exploration of how professors' facilitation is related to literacy teachers' meaning construction process during video-case discussions. *Journal of Reading Education, 39*(1), 15-22.

Arya, P., Christ, T., & Chiu, M. M. (2015). Links between characteristics of collaborative peer video analysis events and literacy teachers' outcomes. *Journal of Technology and Teacher Education, 23*(2), 159-183.

Bartlett, L. (1990). Teacher development through reflective teaching. In Richards, J. C. & Nunan, D. (Eds.), *Second language teacher education* (pp. 202-214). Cambridge University Press.

Boud, D. (2001). Using journal writing to enhance reflective practice. *New Directions for Adult and Continuing Education, 2001*(90), 9–18.

Boud, D., & Walker, D. (2015). Barriers to reflection on experience. In D. Boud, R. Keogh, & D. Walker (Eds.), *Reflection: Turning experience into learning.* Routledge.

Bradbury-Jones, C., Taylor, J., & Herber, O. R. (2014). Vignette development and administration: A framework for protecting research participants. *International Journal of Social Research Methodology, 17*(4), 427–440. https://doi.org/10.1080/13645579.2012.750833

Calandra, B., Sun, Y., & Puvirajah, A. (2014). A new perspective on teachers' video aided reflection. *Journal of Digital Learning in Teacher Education, 30*(3), 104-109. https://doi.org/10.1080/21532974.2014.891880

Christ, T., Arya, P., & Chiu, M. M. (2012). Collaborative peer video analysis: Insights about literacy assessment & instruction. *Journal of Literacy Research, 44*(2), 171-199. https://doi.org/10.1177/1086296X12440429

Craig, C. J. (2013). Opportunities and challenges in representing narrative inquiries digitally. *Teachers College Record, 115*, 1-45. https://doi.org/10.1177/016146811311500405

Davis, E. A. (2006). Characterizing productive reflection among pre-service elementary teachers: Seeing what matters. *Teaching and Teacher Education, 22*, 281-301. https://doi.org/10.1016/j.tate.2005.11.005

Dewey, J. (1933). *How we think: A restatement of the relation of reflective thinking to the educative process.* D.C. Heath & Company.

Eröz-Tuga, P. B. (2013). Reflective feedback session using video recordings. *English Language Teaching, 67*(2), 175-183. https://doi.org/10.1093/elt/ccs081

Friesen, M., & Kuntze, S. (2016). Teacher students analyze texts, comics and video-based classroom vignettes regarding the use of representations – Does format matter? In Proceedings of the 40th Psychology of Mathematics Education Conference, edited by c. Csíkos, A. Rausch and J. Szitány, Vol. 2, 259–266. Szeged: PME.

Friesen, M., & Kuntze, S. (2018). Competence assessment with representations of practice in text, comic and video format. In O. Buchbinder, and S. Kuntze (Eds.), *Mathematics teachers engaging with representations of practice* (pp. 113–130). ICME-14 Monographs. Cham: Springer.

Fund, Z., Court, D., & Kramarski, B. (2002). Construction and application of an evaluative tool to assess reflection in teacher-training courses. *Assessment & Evaluation in Higher Education, 27*(6), 485-499. https://doi.org/10.1080/0260293022000020264

Gibbs, G. (1988). *Learning by doing: A guide to teaching and learning methods*. Further Education Unit.

Henderson, S. J., Horton, R. A., Saito, P. K., & Shorter-Gooden, K. (2016). Validation of assessment vignettes and scoring rubric of multicultural and international competency in faculty teaching. *Multicultural Learning and Teaching, 11*(1), 53–81. https://doi.org/10.1515/mlt-2014-0002

Hughes, R., & Huby, M. (2004). The construction and interpretation of vignettes in social research. *Social Work and Social Sciences Review, 11*(1), 36–51. https://doi.org/10.1921/swssr.v11i1.428

Ibrahim-Didi, K. (2015). Immersion within 360 video settings: Capitalizing on embodied perspectives to develop reflection-in-action within pre-service teacher education. In T. Thomas, E. Levin, P. Dawson, K. Fraser, & R. Hadgraft (Eds.), *Research and development in higher education: Learning for life and work in a complex world* (pp. 235-245). HERDSA.

Jacobs, M., Vakalisa, N. C. G., & Gawe, N. (2011). *Teaching-learning dynamics*. Pearson.

Jasper, M. A. (2003). Nurses' perceptions of the value of written reflection. *Nurse Education Today, 19*(6), 452-463. https://doi.org/10.1054/nedt.1999.0328

Jay, J. K., & Johnson, K. L. (2002). Capturing complexity: A typology of reflective practice for teacher education. *Teaching and Teacher Education, 18*(1), 73–85. https://doi.org/10.1016/S0742-051X(01)00051-8

Jeffries, C., & Maeder, D. W. (2005). Using vignettes to build and assess teacher understanding of instructional strategies. *The Professional Educator, 27*(1&2), 17–28.

Jeffries, C., & Maeder, D. W. (2011). Comparing vignette instruction and assessment tasks to classroom observations and reflections. *The Teacher Educator, 46*(2), 161–175. https://doi.org/10.1080/08878730.2011.552667

Kayapinar, U., & Erkus, A. (2009). Measuring teacher reflection: Development of TRS. *Eurasian Journal of Educational Research, 37*, 144–158.

Kharlay, O., Wei, W., & Philips, J. (2022). How do I teach? Exploring knowledge of reflective practice among in-service EFL teachers in Ukraine. *Teachers and Teaching, 28*(2), 188-205. https://doi.org/10.1080/13540602.2022.2062709

Larrivee, B. (2008). Development of a tool to assess teachers' level of reflective practice. *Reflective Practice, 9*(3), 341–360. https://doi.org/10.1080/14623940802207451

Mathew, P., Prasanth Mathew, P., & Peechattu, P. J. (2017). Reflective practices: A means to teacher development. *Asia Pacific Journal of Contemporary Education and Communication Technology (APJCECT), 3*(1), 126-131.

McGarr, O., & Gallchóir, C. Ó. (2020). Exploring pre-service teachers' justifications for one-to-one technology use in schools: Implications for initial teacher education. *Technology, Pedagogy and Education, 29*(4), 477-490.

Moon, J. A. (2004). *A handbook of reflective and experiential learning: Theory and practice*. Routledge.

Nind, M., & Pepin, B. (2009). Transformation and regulation in educational research. *International Journal of Research & Method in Education, 32*(2), 135–136. https://doi.org/10.1080/17437270902954656

Osterman, K. F., & Kottkamp, R. B. (2015). *Reflective practice for educators: Professional development to improve student learning (2nd ed.).* Corwin Press.

Schon, D. (1983). *The reflective practitioner: How professionals think in action*. Basic Books.

Shanahan, L. E., & Tochelli, A. L. (2014). Examining the use of video study groups for developing literacy pedagogical content knowledge for critical elements of strategy instruction with elementary teachers. *Literacy Research and Instruction, 53*(1), 1-24. https://doi.org/10.1080/19388071.2013.827764

Skilling, K., & Stylianides, G. J. (2020). Using vignettes in educational research: A framework for vignette construction. *International Journal of Research & Method in Education, 43*(5), 541-556. https://doi.org/10.1080/1743727X.2019.1704243

Stecher, B., Le, V. N., Hamilton, L., Ryan, G., Robyn, A., & Lockwood, J. R. (2006). Using structured classroom vignettes to measure instructional practices in mathematics. *Educational Evaluation and Policy Analysis, 28*(2), 101–130. https://doi.org/10.3102/01623737028002101

Tasar, M. F. (2006). Probing pre-service teachers' understandings of scientific knowledge by using a vignette in conjunction with a paper and pencil test. *Eurasia Journal of Mathematics, Science and Technology Education, 2*(1), 53–70. https://doi.org/10.12973/ejmste/75438

Torres, S. (2009). Vignette methodology and culture-relevance: Lessons learned through a project on successful aging with Iranian immigrants to Sweden. *Journal of Cross-Cultural Gerontology, 24*(1), 93-114. https://doi.org/10.1007/s10823-009-9095-9

Turhan, B., & Kırkgöz, Y. (2018). Towards becoming critical reflection writers: A case of English language teacher candidates. *Reflective Practice, 19*(6), 749–762. https://doi.org/10.1080/14623943.2018.1539651

Turhan, B., & Kırkgöz, Y. (2023). A critical and collaborative stance towards retrospective reflection in language teacher education. *European Journal of Teacher Education, 46*(2), 222-240. https://doi.org/10.1080/02619768.2021.1917545

Volkmann, M. J. (2000). Integrating field experience and classroom discussions: Vignettes as vehicles for reflection. East Lansing, MI: National Center for Research on Teacher Learning. (ERIC Document Reproduction Service No. ED442638)

Walshe, N., & Driver, P. (2019). Developing reflective trainee teacher practice with 360-degree video. *Teaching and Teacher Education, 78*, 97-105. https://doi.org/10.1016/j.tate.2018.11.009

Wilkerson, T., Kerschen, K., & Shelton, R. (2018). Pre-service teachers' critical connections to effective mathematical teaching practices: An instructional approach using vignettes. *Action in Teacher Education, 40*(4), 358-373. https://doi.org/10.1080/01626620.2018.1512430

Yang, S. H. (2009). Using blogs to enhance critical reflection and community of practice. *Educational Technology and Society, 12*(2), 11-21.

Yılmaz, E., & Akar, H. (2022). Rethinking the assessment of the quality of teacher reflection by validating an innovative vignette-based instrument and an analytic coding scheme. *Asia Pacific Education Review*, 1-18. https://doi.org/10.1007/s12564-022-09773-6

Zeichner, K. M., & Liston, D. P. (2013). *Reflective teaching: An introduction*. Routledge.

Appendix

Cell 1: Subject Matter-Description: "The teacher started today's lesson with the subject of Kitchen Tools in the 5/H class." (Week 1, P7)

Cell 2: Didactic-Description: "I benefited from some tactics that I learned from humanistic teaching approaches like Suggestopedia in order to calm students down." (Week 4, P8)

Cell 3: Personal-Description: "I felt frustrated and helpless because students never listened to me." (Week 7, P9)

Cell 4: Subject Matter-Personal Opinion: "The topic was games and hobbies. He presented concepts such as like, dislike, enjoy, love, hate and so on. However, I think he should have used images." (Week 4, P4)

Cell 5: Didactic-Personal Opinion: "The teacher was teaching Directions in the 3-D class, and for that topic today, she brought a real map to the classrooms. After working on this map, the teacher conducted an interactive activity through the smartboard on which she showed a map again. I realized that the map on the smartboard attracted the attention of students more." (Week 1, P1)

Cell 6: Personal (Content)-Personal Opinion: "But at least now I know how I should not behave in a young learner classroom. Trying to be the only authority did not work as I expected." (Week 8, P6)

Cell 7: Subject Matter-Linking: "The lack of sufficient physical space and equipment for role-play or storytelling activities was a major problem for the teacher. Unfortunately, there are regional inequalities in this regard in the country. We discussed this last week after a micro-teaching session at university." (Week 7, P11)

Cell 8: Didactic-Linking: "There are many Syrian students in our practicum school. Their existence is good for raising awareness of cultural richness in language classes. Both Turkish and Syrian students can learn how to be respectful toward cultural differences. On the other hand, this can create cultural conflicts among students, which in turn can hinder effective teaching from time to time." (Week 2, P3)

Cell 9: Personal (Content)-Linking: No example was detected in the data.

Cell 10: Subject Matter-Critical Bridging: "We have to use the mother tongue when teaching a foreign language to students at this age because cognitively immature students can need to be exposed to content in their mother tongue [The teacher trainee suggests his own way of content design combining contents in the mother tongue and in English] Thus, students can connect previous and new concepts in their minds in a more structured and memorable way." (Week 6, P11)

Cell 11: Didactic-Critical Bridging: No example was detected in the data.

Cell 12: Personal (Content)-Critical Bridging: No example was detected in the data.

Journal of Theoretical Educational Science, 17(1), 100-121, January 2024
Kuramsal Eğitimbilim Dergisi, 17(1), 100-121, Ocak 2024
[Online]: http://dergipark.org.tr/akukeg
DOI number: http://doi.org/10.30831/akukeg.1294954

Examining Pre-Service Science Teachers' Personal and Enacted Pedagogical Content Knowledge About Seasons

Fen Bilimleri Öğretmen Adaylarının Mevsimler Konusundaki Pedagojik Alan Bilgilerinin İncelenmesi

Aygün KILIÇ*

Received: 10 May 2023 **Research Article** **Accepted:** 2 January 2024

ABSTRACT: This study examined pre-service science teachers' pedagogical content knowledge (PCK) about seasons. The framework of this study was based on the Refined Consensus Model (RCM), and pre-service teachers' topic-specific PCK was also discussed in two dimensions: personal PCK (pPCK) and enacted PCK (ePCK), which are included in this model. The pPCK of the pre-service teachers represents their declarative PCK, and their ePCK represents their dynamic PCK. This qualitative descriptive research included 18 pre-service science teachers. In the study, knowledge of pre-service teachers' pPCK was collected in vignette-based individual semi-structured interviews, and data on their competencies regarding their ePCK, which they revealed in a real classroom environment, were collected through field observations and teaching video recordings. The study's qualitative data were evaluated using two separate rubrics developed by the author using the literature. Accordingly, it was determined that the pre-service teachers who participated in the research had limited pPCK and ePCK on the seasons and that their topic-specific pPCK was partially better quality than their topic-specific ePCK. Moreover, pre-service teachers' ePCK was observed to be different from their pPCK, but their ePCK was based on their pPCK. In this context, it was seen that the results of this research are consistent with RCM.

Keywords: Pedagogical content knowledge (PCK), personal PCK, enacted PCK, teacher education.

ÖZ: Bu çalışmada fen bilimleri öğretmen adaylarının mevsimler konusuna ilişkin Pedagojik Alan Bilgileri (PABları) araştırılmıştır. Bu çalışma rafine fikir birliği modeline dayanmaktadır ve öğretmen adaylarının konuya özgü PABları bu modelin içerdiği kişisel PAB ve uygulanan PAB olmak üzere iki şekilde ele alınmıştır. Öğretmen adaylarının kişisel PABları sahip olduğu PABa ilişkin bilgilerini ve uygulanan PABları ise sınıf ortamında uygulanarak ortaya konulan PABa ilişkin yeterliklerini temsil etmektedir. Nitel betimsel metodun kullanıldığı bu araştırmaya 18 fen bilimleri öğretmen adayı katılmıştır. Araştırmada öğretmen adaylarının kişisel PABlarına ilişkin bilgileri vignetteye dayalı bireysel yarı-yapılandırılmış mülakatlar kullanılarak ve gerçek sınıf ortamında uygulayarak ortaya koydukları PABlarına ilişkin yeterlikleri de ders gözlemleri ve sınıf içi ders video kayıtları kullanılarak toplanmıştır. Araştırmada elde edilen nitel veriler, literatürden yararlanarak yazar tarafından geliştirilen iki ayrı rubrik kullanılarak analiz edilmiştir. Buna göre, araştırmaya katılan öğretmen adaylarının mevsimler konusuna ilişkin kişisel ve uygulanan PABlarının sınırlı olduğu ve konuya özgü kişisel PABlarının konuya özgü uygulanan PABlarına göre kısmen daha iyi kalitede olduğu tespit edilmiştir. Ayrıca, öğretmen adaylarının gerçek sınıf ortamında uygulayarak ortaya koyduğu PABlarının mülakatlarda bildirdiği kişisel PABlarından farklı olduğu, fakat uygulanan PABlarının kişisel PABlarına da dayalı olduğu belirlenmiştir. Bu bağlamda, bu araştırma sonuçlarının rafine fikir birliği modeli ile uyumlu olduğu görülmüştür.

Anahtar kelimeler: Pedagojik alan bilgisi (PAB), kişisel PAB, uygulanan PAB, öğretmen eğitimi.

* Asst. Prof. Dr., Munzur University, Tunceli, Türkiye, aygunkilic@munzur.edu.tr, https://orcid.org/0000-0002-0417-2665

Citation Information
Kılıç, A. (2024). Examining pre-service science teachers' personal and enacted pedagogical content knowledge about seasons. *Kuramsal Eğitimbilim Dergisi [Journal of Theoretical Educational Science], 17*(1), 100-121.

ISSN: 1308-1659

In the literature on teacher education, it has been stated that content knowledge and pedagogical knowledge should be addressed not separately or as competing elements but as components that need to be integrated into teacher education programs to train more qualified teachers or pre-service teachers (Shulman, 1986). In other words, it has been stated that knowing the concepts and procedures related to any science topic very well is not enough for a teacher to teach that topic effectively in the classroom environment (Shulman, 1987). In this context, the concept of Pedagogical Content Knowledge (PCK) has come into prominence as a type of knowledge that is formed by blending content knowledge and pedagogical knowledge. PCK provides the basis of many studies on teacher education in recent years (e.g., Carlson & Daehler, 2019; Gess-Newsome, 2015; Kim, 2020; Mavhunga & Rollnick, 2013; Mazibe, 2020), and many science educators have expressed PCK as an important factor that supports the performance of teachers who improve over time and experience (Eames et al., 2011). Many studies have emphasized that PCK is an important type of knowledge that teachers and pre-service teachers should acquire (e.g., Chan & Hume, 2019; Magnusson et al., 1999; van Driel & Abell, 2010). In these studies, it has also been stated that there is a need to measure and determine the knowledge and skills of teachers and pre-service teachers to confirm whether the practices carried out in teacher education programs and professional development activities are effective (Chan et al., 2019; Coetzee et al., 2022). In this context, it is important to examine pre-service teachers' PCK and classroom teaching practices in terms of basic needs in teacher education.

Different researchers have conceptualized PCK, which teachers should have, in different ways (Mazibe et al., 2020), such as declarative PCK and procedural PCK (Schmelzing et al., 2013), PCK-in-action and PCK-on-action (Park & Oliver, 2008), and dynamic PCK (Alonzo & Kim, 2016). For example, in declarative PCK and PCK-on-action, a teacher expresses conceptual knowledge of students' misconceptions about a science topic in sentences (Schmelzing et al., 2013). Procedural PCK, Dynamic PCK, and PCK-in-action are the practical knowledge of a teacher on activities performed during the lesson (Alonzo & Kim, 2016; Schmelzing et al., 2013). The reasons underlying the conceptualization of PCK in such ways are to emphasize the understanding of the relationship between teachers' knowledge and teaching practices (Alonzo & Kim, 2016) and the importance of how teachers translate their PCK into practice (Park & Suh, 2015). Abell (2008) stated that "PCK is not merely the amount of knowledge in a number of component categories, it is also about the quality of that knowledge and how it is put into action" (p. 1410). In this context, the present study aimed to determine the personal PCK (pPCK) that represents the pre-service science teachers' declarative knowledge of a science topic and the enacted PCK (ePCK) that they revealed in the classroom environment. The theoretical framework of this study was based on the Refined Consensus Model (RCM), which includes these two PCK forms (pPCK and ePCK) (Carlson & Daehler, 2019), and detailed information on this model is presented in the section of the theoretical framework.

Because PCK has been defined as a topic-specific and teacher-specific professional knowledge form (Eames et al., 2011; Lee & Luft, 2008), it is predicted that each teacher and/or pre-service teacher will have a different PCK in each topic (Aydın, 2012; Lankford, 2010; Lee & Luft, 2008). Therefore, in the literature, it has been recommended to conduct studies on topic-specific PCK of teachers and/or pre-service teachers on different science

topics (Abell, 2008; Mazibe, 2020; van Driel et al., 2002). Many researchers have examined and emphasized the topic-specific PCK of science teachers and pre-service teachers using different approaches. For instance, Lankford (2010) has examined PCK on diffusion and osmosis; Mavhunga (2014) on chemical equilibrium; Kim (2020) on force and motion; Mazibe et al. (2020) on graphs of motion; Henze et al. (2008) on models of the solar system and the universe; Coetzee et al. (2022) on electromagnetism. Many PCK studies have been conducted on chemistry, physics, and biology; however, the number of studies examining pre-service teachers' PCK on astronomy-oriented science topics (for example, seasons), which is a separate learning area of the elementary school science program, is quite limited. Nonetheless, it is important to evaluate the PCK of elementary school science teachers and/or pre-service teachers on essential astronomy topics encountered in daily life, such as seasons. In addition, it was emphasized in the literature that learning and teaching this science topic is essential for many reasons, such as misconceptions among students about this topic and the difficulties students face in learning this topic (Sneider et al., 2011; Tsai & Chang, 2005). In this regard, it is thought that the present study will contribute to the literature regarding pre-service science teachers' PCK on the topic of seasons.

Theoretical Framework

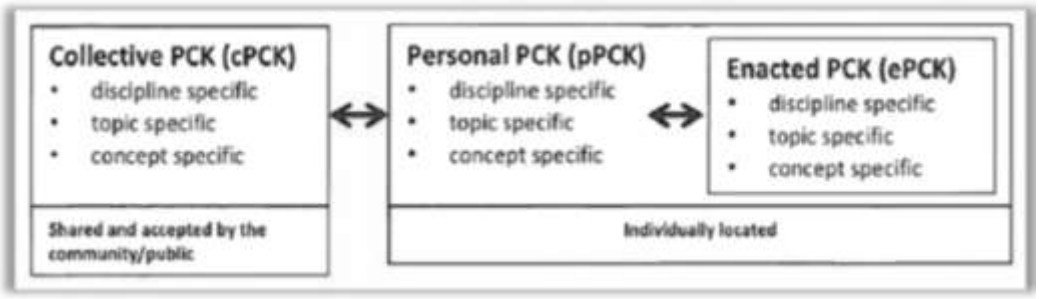
The concept of PCK was first proposed by Shulman (1986) (van Driel et al., 2002). After Shulman (1986), many researchers have introduced various definitions for PCK and its components (e.g., Gess-Newsome, 1999; Magnusson et al., 1999; Mavhunga & Rollnick, 2013; Park & Oliver, 2008; Veal & MaKinster, 1999). For example, Veal and MaKinster (1999) proposed a model, the General Taxonomy of PCK, for future studies on teacher education. In this model, teachers' knowledge is defined at three levels: general (discipline-specific) PCK (e.g., science), domain-specific PCK (e.g., chemistry), and topic-specific PCK (e.g., solubility) (p. 8). Magnusson et al. (1999) defined the concept of PCK with a model of five components: "orientation to teaching science, knowledge of science curricula, knowledge of students' understanding of science, knowledge of instructional strategies, and knowledge of assessment of scientific literacy" (p. 97). Magnusson et al. (1999) explained the topic-specific PCK as (a) curricular knowledge of a specific science topic, (b) knowledge of students' prior knowledge and learning difficulties on a specific science topic, (c) knowledge of instructional strategies and representations used in teaching a specific science topic, and (d) knowledge of assessment strategies and tools used in revealing students' understanding of a specific science topic (Lankford, 2010). In the PCK model of Magnusson et al. (1999), which has been cited the most, it has been stated that "these four elements are universal in the sense that they appear in a variety of general pedagogical models in literature and in teacher education materials" (Barendsen & Henze, 2019, p. 1143). In the present study, pre-service science teachers' PCK on the topic of seasons was examined through these components included in various PCK models in the literature (e.g., Carlson & Daehler, 2019; Gess-Newsome, 2015; Magnusson et al., 1999).

In the last ten years, new models of PCK and its components have been proposed, and many international researchers have organized various PCK summits. In the first PAB summit held in 2012, Gess-Newsome (2015) introduced the Teacher Professional Knowledge and Skill (TPK&S) model with the participants' contributions. This consensus PCK model describes "the overarching role of teacher professional knowledge" (Gess-

Newsome, 2015, p. 30). In the second PCK summit held in 2016, a group of international researchers gathered to examine the differences between different conceptualizations of PCK and review the literature studies on science teachers' PCK (Kim, 2020). As a result of this summit, the Refined Consensus Model (RCM) was revealed (Carlson & Daehler, 2019, p. 83). This latest model recognizes teachers' professional knowledge bases more extensively compared to the TPK&S model and clearly states that subject matter knowledge plays an essential role in teaching (Mazibe, 2020).

Figure 1

Different Components of PCK in the RCM



Note. (Mavhunga, 2019, p. 131)

Figure 1 presents the multi-dimensional nature of PCK in the RCM (Coetzee et al., 2022). Accordingly, three different realms of PCK have been proposed in the RCM: collective PCK (cPCK), personal PCK (pPCK), and enacted PCK (ePCK). The cPCK is "a specialized knowledge base for science teaching that has been articulated and is shared among a group of professionals, which is related to teaching that particular subject matter knowledge to particular students in a particular learning context" (Carlson & Daehler, 2019, p. 88). The pPCK is unique for each teacher and/or pre-service teacher, and it is described as "a specialized form of personal knowledge that includes different knowledge resources related to the teaching and learning of specific science topics" (Alonzo et al., 2019, p. 273). pPCK is different from cPCK (Alonzo et al., 2019); however, cPCK transforms into pPCK when it is amplified and/or filtered by a teacher (Mazibe, 2020). The ePCK is "the specific knowledge and skills utilized by an individual teacher in a particular setting, with a particular student or group of students, with a goal for those students to learn a particular concept, collection of concepts" (Carlson & Daehler, 2019, p. 83). Teachers use the ePCK, which consists of knowledge and skills that guide all aspects of instruction in their classroom teaching practice (Singh, 2021). In other words, ePCK can be expressed as the PCK that is observable during the teaching practices of teachers or pre-service teachers (Aydeniz & Kirbulut, 2014). According to RCM, ePCK is specific to a particular science teaching and does not fully represent teachers' PCK; therefore, it is a subset of pPCK in the model (Carlson & Daehler, 2019). However, it has also been stated that regardless of the level of pPCK of a teacher or pre-service teacher, only some of this existing knowledge informs ePCK in a given situation (Kim, 2020). At the same time, according to RCM, the teaching experience that teachers acquire also shapes their pPCK (Coetzee et al., 2022). This shows that pPCK and ePCK mutually inform each other (Carlson & Daehler, 2019). In RCM, pPCK corresponds to static or declarative PCK, whereas ePCK is associated with dynamic PCK (Chan et al., 2019). In this regard, the

conceptual framework that will guide the current study is the RCM of PCK, and this current study is conceptualized within pPCK and ePCK (Carlson & Daehler, 2019). In this research, pPCK (i.e., declarative PCK) was used as a guide in examining what pre-service science teachers knew or thought about teaching the topic of seasons, and ePCK (i.e., dynamic PCK) was used as a guide to examine their classroom teaching practices.

Literature Review

The pPCK and ePCK in RCM, which are associated with the definition of PCK in the recent literature, are important as they include both the understanding and the enactment of teachers' PCK (Carlson & Daehler, 2019). They also play an essential role in the development of teachers' PCK (Mazibe et al., 2020), as the acquisition and enactment of learned PCK represent two intertwined situations (Park & Oliver, 2008). However, both PCK forms have not yet been researched enough as they have just entered the relevant research literature (e.g., Alonzo & Kim, 2016; Kim, 2020; Mazibe et al., 2020). It is also stated in the literature that pPCK and ePCK should be investigated for reasons such as (1) to explore the knowledge and competencies of teachers in terms of pPCK and ePCK in different science topics (Coetzee et al., 2022), (2) to understand the relationship between teachers' PCK and their teaching practices (Kim, 2020), and (3) to investigate the importance of the gap between pPCK and ePCK (Mazibe et al., 2020). In addition, in the literature, very few studies have been conducted on PCK (i.e., ePCK) captured from classroom teaching observations of teachers and/or pre-service teachers (Chan & Hume, 2019; Gess-Newsome, 2015). Park and Oliver (2008) emphasized that both types of PCK (pPCK and ePCK) are necessary for effective science teaching. In this regard, in the current study, it can be suggested that it is important to examine the knowledge and competencies of pre-service science teachers regarding both PCK types within the scope of a specific science topic.

In most of the studies in the relevant literature, it has been seen that one or multiple PCK components (such as knowledge of students' prior knowledge and learning difficulties and knowledge of instructional strategies and methods) were examined in teachers and/or pre-service teachers (e.g., Belge Can, 2021; Carpendale & Hume, 2019; Chan et al., 2019; Park et al., 2011). However, it has been emphasized that the determination of teachers' topic-specific PCK components is important for a well-developed PCK and effective teaching and that it is necessary to search for the detection of PCK components that teachers have or lack (Akın, 2017; Lee & Luft, 2008). In this context, in the current study, the pre-service science teachers' pPCK and ePCK on the topic of seasons were examined in line with four topic-specific PCK components: Science curriculum, Students' understanding, Instructional strategy and method, and Assessment.

In the literature, there are studies on the description of PCK that teachers and/or pre-service teachers have. In these studies, researchers have used data sources such as lesson plans (Van der Valk & Broekman, 1999), topic-specific PCK tests (Mavhunga & Rollnick, 2013), and content representation (CoRe) (Loughran et al., 2004) to determine teachers' PCK (Chan & Hume, 2019). It is seen that paper and pencil tests have been mostly used to capture teachers' PCK in studies. In most of the recent studies, interviews (Alonzo & Kim, 2016) and lesson observations (Park & Chen, 2012; Park et al., 2011) have also been used in addition to these data sources to capture teachers' PCK. Furthermore, it has been emphasized that it is important to use these two data sources in

the determination of PCK in these studies. However, in the studies, mostly 3-7 pre-service teachers have been interviewed and classroom lessons have been observed (e.g., Park & Oliver, 2008; Sağbilge, 2022). In this case, it can be suggested that the findings of these studies are limited in making a judgment about the levels/quality of PCK of teachers or pre-service teachers. In the current study, individual semi-structured interviews, field observations, and teaching video recordings were used to describe the PCK (pPCK and ePCK) of 18 pre-service science teachers. Accordingly, the vignette was employed in the creation of the interview protocol which was used to determine the pre-service teachers' pPCK (i.e., declarative knowledge). Vignettes are short stories that are developed based on real-life stories (Jeffries & Maeder, 2004) or the findings of previous scientific research (Carlson, 1996) in different formats such as open or closed-ended, fixed or interactive, and in text or video format (Simon & Tierney, 2011). In the literature, it has been stated that the vignette is a good data source to examine the developmental levels (Brovelli et al., 2014) or PCK (McNeill et al., 2015) of teachers and can be used in interviews (Simon & Tierney, 2011). It has also been suggested that the quality of the teachers' PCK would be determined by not only the pPCK they articulate in the interviews but also by how they were put into practice during teaching in the classroom setting (Abell, 2008). In this regard, field observations and teaching video recordings were used to capture pre-service teachers' PCK (i.e., enacted PCK) in the classroom setting. In this direction, it is thought that the current research is important in presenting concrete examples for the studies to be done on the effective development of pre-service teachers' science teaching competencies.

This study aimed to describe the pre-service science teachers' (PSTs') pPCK (i.e., declarative PCK) and ePCK (i.e., dynamic PCK) about the topic of seasons, which is a scientific phenomenon in the curriculum. For this purpose, firstly, the PSTs' knowledge of pPCK and its components was examined. Then, the competencies of PSTs regarding ePCK and its components revealed during teaching practices in the classroom environment were captured. Accordingly, the research questions of this study are as follows:

1. What is the PSTs' knowledge of pPCK and its components involving the topic of seasons?

2. What are the competencies of the PSTs regarding the ePCK and its components involving the topic of seasons?

Method

This qualitative descriptive study examines the pPCK and ePCK of PSTs about the seasons. The purpose of employing this methodology in this study is to provide a direct description of individuals' knowledge and skills regarding an event or circumstance (Lambert & Lambert, 2012; Sandelowski, 2010; Willis et al., 2016). Participants in the study were 18 volunteer PSTs (13 females and 5 males) in their last year of the Science Teacher Education program at the Faculty of Education of a state university in Turkey. These pre-service teachers have never experienced classroom teaching practice in an elementary school until this study.

Data Collection Tools

In this study, interviews, field observations, and teaching video recordings were used as data collection instruments to examine the pPCK and ePCK of PSTs about seasons. pPCKs of PSTs about the seasons were investigated using individual semi-structured

interviews based on vignettes. The interview protocol based on the vignette was developed utilizing relevant literature (e.g., Bailey et al., 2004; Kılıç, 2015; Sneider et al., 2011; Sung & Oh, 2018; Tsai & Chang, 2005) to examine the PSTs' knowledge of the four components of the pPCK in depth. This vignette was created by the researcher from successive stages of the lesson, each containing different events or situations that can occur in a classroom environment during one lesson hour (40 minutes). During the interviews, PSTs were asked open-ended questions about situations such as evaluating teacher-student behaviors and knowledge/opinions in a lesson on seasons based on the vignette, providing alternative suggestions on what the teacher should do in the next lesson, or questioning a specific event presented in the vignette (Schuster et al., 2007). In line with the answers given by PSTs, alternative questions were asked to them when necessary (such as explaining their views with reasons or expressing more clearly why they think the way they do). As a result, it was attempted to investigate PSTs' knowledge of each component of pPCK in depth by allowing PSTs to think more deeply about the concrete situations that can be encountered in the real classroom environment in the vignette and obtain more realistic data. The interview protocol based on this vignette was reviewed by two science educators and edited in line with their suggestions. Pilot interviews were then conducted with two PSTs who did not participate in the study, and the content of the interview protocol was finalized by editing some questions based on the PSTs' responses. All dialogues in the interviews with the PSTs participating in the study were recorded on a voice recorder and then transcribed and analyzed. A part of the vignette provided to PSTs during the interviews in the research is given in Figure 2.

Figure 2

The Vignette on PSTs' Knowledge of Students' Understanding

In the science lesson, teacher Sancar asked his students, "How do you think the seasons are formed?" to determine their prior knowledge of the formation of seasons, and received the following responses:

Student A: Teacher, there are four seasons in a year: spring, summer, fall, and winter… I think the seasons are formed by the Sun revolving around the Earth…

Student B: I think the Earth is the cause of the seasons... The Earth revolves around the Sun, which creates the seasons because the Sun does not move.

Student C: I think that the Earth's rotation around the Sun and its axis creates the seasons…

Student D: Teacher, I think the Sun has a lot to do with how the seasons change... For example, when the Sun is close to the Earth, it is summer and very hot outside. When the Sun moves away from the Earth, it is winter and very cold outside.

Student E: Seasons, I believe, are caused by the tilt of the Earth's axis… Because the Earth's axis is tilted, the distance between the Sun and the Northern and Southern hemispheres changes. This creates seasons…

Student F: Teacher, I think that summer happens on the side of the Earth that faces the Sun, and winter happens on the side that faces away from the Sun... I don't know what causes the other seasons...

After the interviews, the ePCKs of the PSTs were examined about the seasons in their real classrooms while they were teaching. PSTs' ePCKs were analyzed using written notes of field observations and lesson videos. The participant observation method was used in the research process so that the author could directly observe the classroom as it was and keep detailed observation notes on the classroom teaching practice of PSTs. Also, video

recordings of the lessons were used so that the events that happened in the classroom related to the lessons taught by the PSTs could be looked at and described in detail and so that each pre-service teacher's non-verbal actions, such as facial expressions and body movements, could be recorded exactly as they were. The teaching video recordings can also be watched again and again if needed. All of the pre-service teachers who took part in the study and the school administrators who needed to give permission did so. These data collection tools are among the data sources that have been frequently used to determine the PCK of teachers and/or pre-service teachers in recent years (Chan & Hume, 2019). Each PST participating in the study was given one lesson hour (40 minutes) for classroom teaching practices.

Data Analysis

The data were analyzed in a way that was similar to what was done in other studies (e.g., Chan et al., 2019; Mavhunga & Rollnick, 2013; Mazibe, 2020; Park et al., 2011). Since the literature usually uses qualitative data sources to measure the PCK of teachers and/or pre-service teachers, it has become common to use rubrics to analyze qualitative data (Chan et al., 2019). Accordingly, in the present study, the two forms of PCK given in the theoretical framework, pPCK and ePCK, were used as guides and two assessment rubrics were created for both pPCK (Appendix 1) and ePCK (Appendix 2) regarding the seasons, taking into account the four topic-specific components of PCK. These rubrics were designed by the author by examining many scales used in the literature to assess teachers' PCK (e.g., Mazibe, 2020; Mazibe et al., 2020). The answers given to the questions in the interviews used to evaluate the knowledge of PSTs about the pPCKs in relation to the seasons and the competencies of PSTs about the ePCKs in relation to the seasons were evaluated in four categories: limited, basic, developing, and exemplary (Mavhunga & Rollnick, 2013; Mazibe et al., 2020). For this process to be reliable, the items on both evaluation lists were reviewed and changed by a science educator who is an expert on the topic, and the changes were made based on his suggestions. As a result, rubrics with seven items for topic-specific pPCK and seven items for topic-specific ePCK were created to be used in analyzing the data from this study. The transcripts of the vignette-based individual semi-structured interviews were read several times, examined in detail, and evaluated according to the rubric prepared for the pPCK in order to analyze the data obtained as a result of the study. Likewise, the qualitative data obtained from field observations and teaching video recordings during the PSTs' teaching practices in the classroom environment were analyzed in depth and evaluated according to the rubric prepared for ePCK. Furthermore, a science educator re-analyzed the data obtained from the data sources of the three PSTs participating in the study using both rubrics to ensure the reliability of the data obtained (Miles & Huberman, 1994). For pPCK and ePCK, the average agreement between the author and expert analyses was 86%.

Ethical Procedures

Ethical approval and written permission were obtained from the Munzur University Non-Interventional Research Ethics Committee with the decision dated 30.03.2023 and numbered 2023/05-09.

Results

PSTs' Topic-Specific pPCK

The pPCKs of the PSTs participating in this study were presented under separate subheadings specifically within the context of PCK components. PSTs' pPCKs (i.e., declarative PCK) about seasons are given in Table 1.

Table 1

PSTs' Topic-Specific pPCK about Seasons

pPCK		Limited	Basic	Developing	Exemplary
Science curriculum	Concepts and achievements in the science curriculum	11	4	3	-
Students' understanding	Students' prior knowledge	10	4	4	-
	Students' learning difficulties (misconceptions) and their reasons	5	6	7	-
Instructional strategy and method	Instructional strategy and method	12	5	1	-
	Representations	6	10	2	-
Assessment	Assessment approaches and methods	14	4	-	-
	Assessment tools and their use	12	6	-	-

Science curriculum knowledge of PSTs

When the curriculum knowledge about the concepts and achievements of the seasons topic in the science curriculum was examined, it was found that most pre-service teachers did not know the program content. Regarding the concepts and achievements related to this science topic, 11 PSTs made limited explanations, such as students knowing the four seasons (spring, summer, fall, and winter), the Earth, and the Sun. Four PSTs explained these concepts at the basic level by giving an example from the objectives of the science program in addition to these concepts. On the other hand, three PSTs mentioned the concepts and achievements related to the formation of seasons, such as the Earth's motions, the Earth's axis of rotation, axis tilt, and the Earth's orbital motion around the Sun at the developing level.

Students' understanding knowledge of PSTs

In Table 1, when the findings declared by the PSTs in the interview regarding the students' understanding of the formation of seasons are analyzed, some of the PSTs either stated that they did not know the students' prior knowledge on the topic (n=4) or they gave limited answers with statements such as students know the names of the seasons and/or they know the order of the seasons (n=6). Furthermore, four PSTs declared that students have prior knowledge of the Earth's rotation and circulation, the Sun's motion, and the seasons' names at the basic level. According to the other four PSTs, the students had developed prior knowledge about the topic, such as the rotational motion of the Earth around its axis and the results of this motion, the motion of the Sun, and the direction and

motion of the Earth's orbit around the Sun. The knowledge of the PSTs on students' learning difficulties (misconceptions) and their causes regarding the formation of the seasons was attempted to be determined by open-ended questions posed to them during the interview using the vignette. Accordingly, seven PSTs defined two or three misconceptions of the students about this science topic, such as seasons are formed by the change in the distance between the Earth and the Sun or seasons are formed by the rotation of the Sun around the Earth and stated their reasons with developing explanations. While six PSTs were able to identify one or two misconceptions of the students in the vignette presented to them, they were not able to express fully their reasons. On the other hand, the other PSTs both stated that students would not have any difficulty with this science topic and that there would be no misconceptions, and they could not define the misconceptions in the vignette. Also, three of these PSTs were found to have misconceptions about the formation of seasons. Below, you may see some examples of interviews made with PSTs.

> Researcher: (vignette in Figure 2) What do you think about the students' answers to teacher Sancar's question?
>
> PST-8: ...before studying this topic, students should be familiar with the movements of the Sun and the movements of the Earth... They also need to know what happens as a result of these because if students do not learn them well, they will have difficulty learning the next topics, and misconceptions will occur... For example, here, Student A says that the seasons are formed by the Sun revolving around the Earth, which is sad because the child does not know that the Earth revolves around the Sun... This child is in misconception... Student C also said that seasons occur when the Earth rotates both around the Sun and around its axis, which is wrong because seasons occur when the Earth rotates around the Sun... Student B thinks as I say, so this student is right... The other student (Student D) says that the seasons are formed not by the tilt of the Earth's axis but by the proximity or distance of the Earth to the Sun, which is wrong... So, Student D has incomplete knowledge, and I think she/he does not know the movements of the Earth and the Sun... Student F is also wrong... Only Student E is exactly right because he knows that the seasons are formed by the tilt of the Earth's axis... Other students already have learning difficulties because they do not have this information...
>
> PST-1: Student A thinks that the Earth does not revolve around the Sun, but the Sun revolves around the Earth, this student has the misconception... That is, he did not say that seasons are formed by revolving the Earth around the Sun... Student C has incomplete knowledge, I think this student confused the formation of day and night with the formation of seasons... Student D knows something because he said that the Sun plays a role in the formation of the seasons, but the Sun is not the only factor in the formation of the seasons... Here, the student did not know about the tilt of the axis, and he fell into a misconception... Here, she/he made a mistake in the proximity and distance of the Earth to the Sun, that is, she/he thought that summer occurs when the Earth is close, and winter occurs when the Earth is far away... Student E said exactly the right thing, that is, this student does not have a misconception... She/He said that seasons are formed due to the tilt of the Earth's axis... I believe she/he has the right prior knowledge... Student F is also in misconception...

Instructional strategy and method knowledge of PSTs

When the findings in Table 1 regarding the teaching strategies and methods declared by the PSTs in the interview were examined, it was observed that 12 PSTs would use teacher-centered teaching strategies in teaching the topic of seasons either by making plain explanations on the topic or by using only the question-answer technique. Furthermore, these PSTs stated that they did not know how to plan and teach their lesson or how to create an activity because their knowledge of the content of this science topic was insufficient. However, five PSTs stated their explanations on how to create learning activities, such as showing the Earth's orbit around the Sun with a model for the formation

of seasons, etc., by using hands-on materials (e.g., globes) in a way that students would be active at the basic level. Only one PST gave developing explanations on how to create student-centered learning environments based on the constructivist learning approach and making students active in the process of teaching this topic. It was observed that six of the PSTs who participated in the study were at the limited level because they either stated that they would not use any representation while teaching the topic of the seasons or that they would use a representation to inform students about this topic (i.e., in a teacher-centered way). Furthermore, 10 PSTs gave basic explanations on how they would use representations such as illustrations, videos, simulations, hands-on materials, models, etc. either to summarize the lesson or to reinforce students' knowledge on this topic. Two PSTs stated that they would use representations to concretize/visualize this scientific phenomenon or eliminate students' misconceptions about it with developing-level explanations. They also said that they would use these representations to motivate, encourage, and engage students in science lessons.

Assessment knowledge of PSTs

When the findings reported by the PSTs in the interview regarding the assessment component are analyzed in Table 1, 14 PSTs made limited explanations by stating that they would evaluate students' understanding of the seasons at the end of the lesson (i.e., the traditional assessment approach). Other PSTs provided basic-level explanations. Regarding assessment tools and their use, 12 PSTs stated that they would assess students' learning on this topic at the end of the lesson through question-answer (oral roll call), true-false or multiple-choice tests, or homework. Furthermore, six PSTs stated that they would assess students' understanding of how the seasons are formed at the basic level by making drawings, assigning project homework, or evaluating students' understanding through student work.

PSTs' Topic-Specific ePCK

The ePCKs of the PSTs participating in the study were presented under separate subheadings specifically within the context of PCK components. PSTs' ePCKs (i.e., dynamic PCK) about seasons are given in Table 2.

Table 2

PSTs' Topic-Specific ePCK about Seasons

ePCK		Limited	Basic	Developing	Exemplary
Science curriculum	Referring to the concepts and achievements in the science curriculum	9	6	3	-
Students' understanding	Questioning students' prior knowledge and learning difficulties (misconceptions)	11	5	2	-
	Recognizing and acknowledging students' prior knowledge and learning difficulties (misconceptions)	13	4	1	-
Instructional strategy and method	Use of instructional strategy and method	11	5	2	-
	Use of representations	10	6	2	-
Assessment	Assessment approaches and methods	14	4	-	-
	Use of assessment tools	14	4	-	-

Science curriculum knowledge of PSTs

When the science curriculum component was examined, it was observed that only three PSTs referred to concepts such as the position of the Earth, the motion of the Earth's orbit around the Sun and its consequences, and the tilt of the Earth's axis of rotation (axial tilt) about the seasons, and they taught their lessons by making logical connections between most of these concepts. Six PSTs were identified as being at the basic level because they focused on some of the concepts about this topic (such as the structure of the Sun, the sizes of the Earth and the Sun, and the rotational and orbital motions of the Earth) and only taught by establishing relationships between these concepts. Nine PSTs were found to be at a limited level because they did not teach most of the important concepts about how seasons come to be; they mostly focused on concepts that had nothing to do with each other, and they did not make connections between concepts related to this topic.

Students' understanding knowledge of PSTs

When the results of the ePCK of the PSTs on students' understanding are examined in Table 2, 11 PSTs were assessed to be at the limited level because they did not question students' possible prior knowledge and learning difficulties (misconceptions) about seasons, and five PSTs were assessed to be at the basic level because they partly questioned them by asking only a few. It was observed that the other PSTs identified students' prior knowledge and at least two misconceptions about the topic, such as the summer and winter seasons are formed by the side of the Earth facing the Sun. For example, these PSTs questioned students' prior knowledge and misconceptions by asking them some open-ended questions about the topic, by having a few students draw on the board how the seasons occur, or by asking questions based on pictures or a model of the Solar system. Hence, the competencies revealed by these PSTs during the lesson were scored as developing. During the lesson, 11 out of 13 PSTs did not question students' prior knowledge and learning difficulties regarding the seasons and, therefore, did not teach their

lessons by taking learners' thinking into account. The two PSTs partially questioned the students by asking a few questions, but it was observed that they did not take the students' answers into account and did not recognize their misconceptions. Therefore, the competencies enacted by these PSTs in a real classroom environment were scored as limited. Four PSTs questioned students' prior knowledge and learning difficulties but continued the lesson by partially taking into account the students' thoughts and, therefore, were scored as basic. The other PST was scored as developing because she/he recognized and acknowledged students' prior knowledge, such as that there are four seasons, that the Earth moves in a circle, and that the Sun is bigger than the Earth, and some of the students' misconceptions about the seasons during the lesson.

Instructional strategy and method knowledge of PSTs

During the classroom teaching practices, it was observed that 11 PSTs used the lecture method while teaching the topic of the seasons and did some of the teacher-centered activities in which students were passive. It was also observed that these PSTs were weak in classroom management. When the findings related to the use of instructional strategy and method in Table 2 are examined, it is observed that the competencies revealed by these PSTs during the lesson were scored as limited. The competencies of five PSTs during the lesson were scored as basic because they explained the topic by drawing shapes on the board or asking questions through some representations such as pictures, slides, videos, and models/mock-ups and then made explanations about the topic and were partially successful in classroom management. On the other hand, two PSTs used a model/mock-up of the Earth and a flashlight to have students explore the topic by asking questions, or they used a model/mock-up of the Solar system and asked questions related to the topic and created discussion environments primarily with dialogues between the teacher and the students, or they showed a video or simulation and asked various questions about the formation of the seasons and made explanations. In this regard, the competencies of these PSTs during the lesson were also scored as developing. 10 PSTs who participated in the study were scored as limited because they either did not use any representation about the topic or used representations such as illustrations and diagrams to make direct explanations about the seasons. The competencies of 6 PSTs were scored as basic during the lesson because they partially used representations such as illustrations and models/mock-ups to summarize the related topic at the end of the lesson or to question students' prior knowledge about the seasons, etc. The other PSTs were rated at the developing level for their competencies in the use of representations such as hands-on materials (e.g., globes, flashlights), models, videos, and simulations during classroom teaching practices.

Assessment knowledge of PSTs

Fourteen of the PSTs participating in the study either did not assess students' understanding of the seasons or tried to assess it by asking questions to a few students at the end of the lesson and did not use any assessment tool. Therefore, the competencies of these PSTs regarding assessment approaches and methods and the use of assessment tools during the lesson were scored as limited (Table 2). The four PSTs tried to partially evaluate the students' understanding of the topic from the beginning to the end of the lesson and had them draw and explain how the seasons are formed on a piece of paper and then collected

the papers. The competencies that these PSTs demonstrated in a real classroom environment were scored as basic. None of the PSTs' competencies in assessment approaches and methods and use of assessment tools demonstrated during the lesson were rated as developing or exemplary.

Discussion and Conclusion

This study examined the pPCK (i.e., declarative PCK) and ePCK (i.e., dynamic PCK) of the PSTs on the topic of seasons. According to the data obtained from the vignette-based individual semi-structured interviews held with the PSTs who participated in the research, the pre-service teachers had a limited level of pPCK on the topic of seasons. According to the data obtained as a result of the examinations made during the lectures of the PSTs in the real classroom environment, the pre-service teachers had a limited level of ePCK related to the topic of seasons. The PSTs' knowledge of all components of topic-specific pPCK and the PSTs' competencies regarding all components of topic-specific ePCK during the teaching practice were mostly limited. Regarding the other components of the topic-specific pPCK, the majority of the PSTs' knowledge of students' understanding was the best, and their knowledge of assessment was the worst. Regarding the other components of the topic-specific ePCK, the knowledge of the science curriculum was the best, and the knowledge of assessment was the worst. Likewise, in previous studies, it has been indicated that the knowledge and competencies of teachers and/or pre-service teachers on the components of PCK would not develop at the same level (Aydın, 2012; Barendsen & Henze, 2019). In studies, it has also been observed that the PSTs' knowledge of assessment and their competencies regarding assessment revealed during the lecture were quite weak (Aydın, 2012; Sağbilge, 2022). The general reasons for this situation may include the fact that pre-service science teachers do not take applied courses on assessment approaches and methods and assessment tools during their university education, that they do not attach importance to the process of evaluating what students have learned during their lesson planning and teaching practices as well as students' learning experience, etc. Furthermore, in the present study, none of the pre-service teachers' explanations regarding the components of pPCK and the competencies regarding the components of ePCK during the lecture were at an exemplary level. Similarly, in the literature, it has been stated that teachers' and/or pre-service teachers' knowledge of topic-specific pPCK in various science topics and their competencies regarding topic-specific ePCK were weak (e.g., Barendsen & Henze, 2019; Mavhunga, 2014). The reasons for this may be, in particular, pre-service teachers' lack of content knowledge and teaching experience (Barendsen & Henze, 2019; Kim, 2020; Kind, 2009; Mazibe, 2020; Rollnick et al., 2008). Many PSTs who participated in the current study also admitted during the interviews that they lacked content knowledge on the relevant science topic. In the literature, it has been emphasized that teachers' content knowledge is a prerequisite for PCK (Kim, 2020; Mazibe, 2020). However, in some studies, it has been indicated that a good level of content knowledge alone is not sufficient for teachers or pre-service teachers to have a strong PCK (Davidowitz & Potgieter, 2016; Rollnick & Mavhunga, 2014; van Driel et al., 1998). In this context, it has been stated that in addition to content knowledge, teachers and/or pre-service teachers need to gain experience in classroom teaching practices, attend education-related seminars, and exchange opinions with more experienced teachers to develop their topic-specific PCK (Aydın, 2012; Henze et al., 2008; van Driel et al., 1998). In short, it is suggested that PSTs should first develop

their cPCK regarding their knowledge and competencies in the teaching profession during their undergraduate education. cPCK, which is defined as an expert knowledge base for science teaching in RCM (Carlson & Daehler 2019), also includes teachers' pPCK and ePCK. According to this model, to what extent this form of PCK is amplified and/or filtered by each teacher or pre-service teacher, they can develop their pPCK and ePCK proportionally (Carlson & Daehler 2019; Mazibe, 2020). In this respect, it is suggested that educational activities should be carried out through in-service training for teachers or teacher education programs for pre-service teachers to share, learn, and gain knowledge and skills about cPCK.

According to the qualitative findings obtained in this study, the PSTs' pPCK on the topic of seasons that they articulated in the interviews and their competencies regarding ePCK that they revealed in the classroom environment mostly differed. However, it can also be said that pre-service teachers' competencies they enacted during the teaching practices (ePCK) were heavily dependent on the available knowledge (pPCK) they articulated in vignette-based interviews, but their pPCK did not fully reflect their ePCK captured in the real classroom environment. This may be explained by the fact that the competencies revealed by the PSTs in their use of topic-related representations during the teaching practice (ePCK) relied on their knowledge of instructional strategies and methods or representations that PSTs declared in the interviews; however, PSTs did not apply their knowledge of representations in the classroom environment. This result is consistent with RCM, which positions ePCK as a subset of pPCK (Carlson & Daehler 2019) (Figure 1). In addition, PSTs' components of topic-specific pPCK and ePCK were mostly determined to differ. For example, when the data on the students' understanding component of PCK was examined, it was seen that the PSTs did not recognize and acknowledge the students' misconceptions on this topic in teaching practices or did not teach the topic by considering students' thoughts, although some of them articulated the students' explanations of learning difficulties (misconceptions) and the causes of these difficulties regarding the formation of seasons at a developing level in the interview (see Tables 1 and 2). The findings of this research support the relationships between pPCK and ePCK specified in the RCM (Carlson & Daehler, 2019). According to this model, PSTs' pPCK forms the basis for the ePCK they reveal during classroom teaching practices and serves as an existing knowledge and skill repository from which they can benefit (Alonzo et al., 2019). At the same time, according to RCM, PSTs' pPCK and their ePCK are not expected to be compatible since ePCK is flexible and changes for a particular learning environment, a particular student group, a particular science topic, or is formed and used at that moment in the classroom setting (Alonzo et al., 2019). The results of some studies (Kim, 2020; Mazibe, 2020; Mazlum Güven & Yiğit, 2020) on different science topics in the literature are consistent with the results of this research.

Considering the study's overall findings, it was seen that most of the PSTs' topic-specific pPCK was partially of better quality than their topic-specific ePCK. For instance, some pre-service teachers partially declared their knowledge and skills about the students' prior knowledge of seasons and how to determine them in a theoretical way, but they were weak in applying their knowledge and skills while teaching the lessons on the relevant topic in a real classroom environment. It can be suggested that the general reasons for this are that the general or topic-specific pedagogical knowledge and competencies of PSTs were weak, that very few lessons are directly related to discipline-specific (science-

specific) pedagogical knowledge and skills in the science teacher education program, and that most importantly, very few opportunities regarding teaching practices are provided to PSTs in a real classroom environment during undergraduate education (Alonzo & Kim, 2016; Mazibe et al., 2020). From these perspectives, it was concluded that pre-service teachers need to carry out more instructional practices in order to transfer the PCKs (pPCK) they have or acquired to their lessons in the real classroom environment. Moreover, it is suggested that pre-service teachers should constantly reflect on the experiences they acquired during these teaching practices or different/initial experiences of themselves (reflective practice) (Park & Oliver, 2008). When the findings regarding the topic-specific pPCK and ePCK components of PSTs were examined, it was seen that the PSTs' pPCK was of slightly better quality than their ePCK in all components except the science curriculum component. This may be explained by the fact that even though pre-service teachers consider the students' misconceptions on the relevant topic and effectively plan the instructional activities and the assessment process, they may be weak while practicing the designed lesson plan in the real classroom environment. Compared to the other components, the general reason why the PSTs' competencies of the topic-specific science curriculum component in the classroom environment (ePCKs) were of slightly better quality than their knowledge of this component (pPCKs) might be that pre-service teachers repeat their content knowledge on the relevant topic during the planning of the lesson and review the concepts, achievements, etc. in the science program before teaching the lesson. Likewise, in previous studies, it has been stated that the declarative knowledge and teaching practices of teachers and/or pre-service teachers would develop differently (Alonzo & Kim, 2016; Mazibe, 2017; 2020; Mazlum Güven & Yiğit, 2020). In this respect, the results of this study are thought to be important in terms of presenting empirical evidence while explaining the relationships between teachers' knowledge and their classroom teaching practices.

Previous studies in the literature have determined that the quality of topic-specific PCK of teachers/pre-service teachers differs based on the topic. For example, Aydın (2012) has examined the PCKs of two chemistry teachers on electrochemistry and radioactivity and has stated that there are two different types of PCK for two different science topics: PCK A for teaching electrochemistry and PCK B for teaching radioactivity. Therefore, as in the current study, it was concluded that it would be more appropriate to research PCK in its topic-specific dimension since the quality of pre-service teachers' pPCK and ePCK may vary within the scope of different science topics (Mazibe, 2020; Rollnick & Mavhunga, 2014). Accordingly, it can be recommended to conduct studies examining PCK and classroom teaching practices of science teachers/pre-service teachers regarding different science topics that students have difficulty learning in order to provide concrete evidence to researchers on teacher education. In addition, in this study, the vignette was used to examine the topic-specific pPCK of the PSTs, and vignette-based individual semi-structured interviews within the scope of the related science topic were held with the PSTs. Based on the current research findings, it is thought that the vignettes used during the interviews in the study have a significant potential to provide valid and reliable information about the quality of topic-specific pPCK of pre-service teachers. It can be recommended the vignette be used in PCK studies to be carried out in the field of teacher education. It is suggested that presenting concrete situations to participants through written short stories or short videos is more effective in terms of assessing these concrete cases and

expressing views. Therefore, it is thought that it is important for researchers to explore or question the knowledge and skills of participants in different dimensions.

Conflicts of Interest

The author declares that there is no conflict of interest.

Author Bio:

Aygün Kılıç is an assistant professor at the Tunceli Vocational Higher School City, Munzur University. She did her PhD on science education. Her research interests are teacher education, PCK, TPACK, blended learning, technology use in science teaching, instructional technologies, and reflective practices.

References

Abell, S. (2008). Twenty years later: Does pedagogical content knowledge remain a useful idea? *International Journal of Science Education*, *30*, 1405-1416.

Akın, F. N. (2017). *The nature of interplay among components of pedagogical content knowledge in reaction rate and chemical equilibrium topics of novice and experienced chemistry teachers* [Unpublished doctoral dissertation]. Middle East Technical University.

Alonzo, A. C., Berry, A., & Nilsson, P. (2019). Unpacking the complexity of science teachers' PCK in action: Enacted and personal PCK. In A. Hume, R. Cooper, & A. Borowski (Eds.), *Repositioning Pedagogical Content Knowledge in Teachers' Knowledge for Teaching Science* (pp. 271-286). Springer.

Alonzo, A. C., & Kim, J. (2016). Declarative and dynamic pedagogical content knowledge as elicited through two video-based interview methods. *Journal of Research in Science Teaching*, *53*(8), 1259-1286.

Aydeniz, M., & Kirbulut, Z. D. (2014). Exploring challenges of assessing pre-service science teachers' pedagogical content knowledge (PCK). *Asia-Pacific Journal of Teacher Education*, *42*(2), 147-166.

Aydın, S. (2012). *Examination of chemistry teachers' topic-specific nature of pedagogical content knowledge in electrochemistry and radioactivity* [Unpublished doctoral dissertation]. Middle East Technical University.

Bailey, J. M., Prather, E. E., & Slater, T. F. (2004). Reflecting on the history of astronomy education research to plan for the future. *Advances in Space Research*, *34*(10), 2136-2144.

Barendsen, E., & Henze, I. (2019). Relating teacher PCK and teacher practice using classroom observation. *Research in Science Education*, *49*, 1141-1175.

Belge Can, H. (2021). Enacted pedagogical content knowledge profiles of chemistry teachers. *Journal of Educational Issues*, *7*(1), 565-583.

Brovelli, D., Bölsterli, K., Rehm, M., & Wilhelm, M. (2014). Using vignette testing to measure student science teachers' professional competencies. *American Journal of Educational Research*, *2*(7), 555-558.

Carlson, B. E. (1996). Dating violence: Student beliefs about consequences. *Journal of Interpersonal Violence*, *11*, 3-18.

Carlson, J., & Daehler, K. R. (2019). The refined consensus model of pedagogical content knowledge in science education. *In Repositioning Pedagogical Content Knowledge in Teachers' Knowledge for Teaching Science* (pp. 77–92). Springer.

Carpendale, J., & Hume, A. (2019). Investigating practising science teachers' pPCK and ePCK development as a result of collaborative CoRe design. *In Repositioning Pedagogical Content Knowledge in Teachers' Knowledge for Teaching Science* (pp. 223-250). Springer.

Chan, K. K. H., & Hume, A. (2019). Towards a consensus model: Literature review of how science teachers' pedagogical content knowledge is investigated in empirical studies. In A. Hume, R. Cooper, & A. Borowski (Eds.), *Repositioning Pedagogical Content Knowledge in Teachers' Knowledge for Teaching Science* (pp. 3-76). Springer.

Chan, K. K. H., Rollnick, M., & Gess-Newsome, J. (2019). A grand rubric for measuring science teachers' pedagogical content knowledge. In A. Hume, R. Cooper, & A. Borowski (Eds.), *Repositioning Pedagogical Content Knowledge in Teachers' Knowledge for Teaching Science* (pp. 251-269). Springer.

Coetzee, C., Rollnick, M., & Gaigher, E. (2022). Teaching electromagnetism for the first time: A case study of pre-service science teachers' enacted pedagogical content knowledge. *Research in Science Education*, *52*, 357-378.

Davidowitz, B., & Potgieter, M. (2016). Use of the Rasch measurement model to explore the relationship between content knowledge and topic-specific pedagogical content knowledge for organic chemistry. *International Journal of Science Education*, *38*(9), 1483-1503.

Eames, C. W., Williams, P. J., Hume, A. C., & Lockley, J. (2011). CoRe: A way to build pedagogical content knowledge for beginning teachers. *Wellington: Teaching and Learning Research Initiative.* Retrieved from http://researchcommons.waikato.ac.nz/handle/10289/7399

Gess-Newsome, J. (1999). Pedagogical content knowledge: an introduction and orientation. In J. Gess-Newsome & N. G. Lederman (Eds.), *Examining Pedagogical Content Knowledge* (pp. 3-17). Kluwer.

Gess-Newsome, J. (2015). A model of teacher professional knowledge and skill including PCK. In: A. Berry, P. Friedrichsen, & J. Loughran (Eds.), *Re-Examining Pedagogical Content Knowledge in Science Education* (pp. 28-42). Routledge.

Henze, I., van Driel, J. H., & Verloop, N. (2008). Development of experienced science teachers' pedagogical content knowledge of models of the solar system and the universe. *International Journal of Science Education*, *30*(10), 1321-1342.

Jeffries, C. M., & Maeder, D. W. (2004). Using vignettes to build and assess teacher understanding of instructional strategies. *Professional Educator*, *27*(1/2), 17-28.

Kılıç, A. (2015). *The effect of technological pedagogical content knowledge (TPACK) based blended learning environment on pre-service science teachers' TPACK and classroom practices involving the topics of basic astronomy* [Unpublished doctoral dissertation]. Firat University.

Kim, J. (2020). *Connecting pedagogical content knowledge (PCK) to teaching practice: Investigating physics teachers' enacted PCK, personal PCK, and engagement with student ideas in classroom discussions* [Unpublished doctoral dissertation]. Michigan State University.

Kind, V. (2009). Pedagogical content knowledge in science education: perspectives and potential for progress. *Studies in Science Education, 45*(2), 169-204.

Lambert, V. A., & Lambert, C. E. (2012). Qualitative descriptive research: an acceptable design. *Pacific Rim International Journal of Nursing Research*, *16*, 255-256.

Lankford, D. (2010). *Examining the pedagogical content knowledge and practice of experienced secondary biology teachers for teaching diffusion and osmosis* [Unpublished doctoral dissertation]. University of Missouri.

Lee, E., & Luft, J. A. (2008). Experienced secondary science teachers' representation of pedagogical content knowledge. *International Journal of Science Education*, *30*, 1343-1363.

Loughran, J., Mulhall, P., & Berry, A. (2004). In search of pedagogical content knowledge in science: Developing ways of articulating and documenting professional practice. *Journal of Research in Science Teaching*, *41*(4), 370-391.

Magnusson, S., Borko, H., & Krajcik, J. (1999). Nature, sources, and development of pedagogical content knowledge for science teaching. In J. Gess-Newsome & N. Lederman (Eds.), *Examining Pedagogical Content Knowledge: The Construct and Its Implications for Science Education* (pp. 95–132). Kluwer Press.

Mavhunga, E. (2014). Improving PCK and CK in pre-service chemistry teachers. In H. Venkat, M. Rollnick, M. Askew & J. Loughran (Eds.), *Exploring mathematics and science teachers' knowledge: Windows into teacher thinking*, (pp. 31-48). Oxford: Routledge. Retrieved from https://www.researchgate.net/publication/271827268_Improving_PCK_and_CK_in_chemistry_pre-service_teachers/link/5714bb5c08ae8b02e65de1a2/download

Mavhunga, E. (2019). Exposing pathways for developing teacher pedagogical content knowledge at the topic level in science. *Repositioning Pedagogical Content Knowledge in Teachers' Knowledge for Teaching Science* (pp. 129-148). Springer.

Mavhunga, E., & Rollnick, M. (2013). Improving PCK of chemical equilibrium in pre-service teachers. *African Journal of Research in Mathematics, Science and Technology Education*, *17*(1-2), 113-125.

Mazibe, E. N. (2017). *Teaching graphs of motion: Translating pedagogical content knowledge into practice* [Unpublished master's thesis]. University of Pretoria.

Mazibe, E. N. (2020). *The relationship between teachers' pedagogical content knowledge about electrostatics and learners' performance* [Unpublished doctoral dissertation]. University of Pretoria.

Mazibe, E. N., Coetzee, C., & Gaigher, E. (2020). A comparison between reported and enacted pedagogical content knowledge (PCK) about graphs of motion. *Research in Science Education*, *50*(3), 941-964.

Mazlum Güven, E., & Yiğit, N. (2020). Investigating pre-service science teachers' classroom performance in the context of pedagogic content knowledge. *Journal of Abant İzzet Baysal University Education Faculty*, *20*(3), 1590-1607.

McNeill, K. L., González-Howard, M., Katsh-Singer, R., & Loper, S. (2015). Pedagogical content knowledge of argumentation: Using classroom contexts to assess high-quality PCK rather than pseudoargumentation. *Journal of Research in Science Teaching*, *53*(2), 261-290.

Miles, M. B., & Huberman, A. M. (1994). *Qualitative data analysis: An expanded sourcebook*. Sage.

Park, S., & Chen, Y.-C. (2012). Mapping out the integration of the components of pedagogical content knowledge (PCK): examples from high school biology classrooms. *Journal of Research in Science Teaching*, *49*(7), 922-941.

Park, S., Jang, J. Y., Chen, Y. C., & Jung, J. (2011). Is pedagogical content knowledge (PCK) necessary for reformed science teaching?: Evidence from an empirical study. *Research in Science Education*, *41*, 245-260.

Park, S., & Oliver, S. T. (2008). Revisiting the conceptualisation of pedagogical content knowledge (PCK): PCK as a conceptual tool to understand teachers as professionals. *Research in Science Education*, *38*, 261-284.

Park, S., & Suh, J. (2015). Trajectory from portraying toward assessing PCK: Drives, dilemmas, and directions for future research. In A. Berry, P. Friedrichsen, & J. Loughran (Eds.), *Re-examining pedagogical content knowledge in science education*, (pp. 104–119). Routledge.

Rollnick, M., Bennett, J., Rhemtula, M., Dharsey, N., & Ndlovu, T. (2008). The place of subject matter knowledge in pedagogical content knowledge: A case study of South African teachers teaching the amount of substance and chemical equilibrium. *International Journal of Science Education*, *30*(10), 1365-1387.

Rollnick, M., & Mavhunga, E. (2014). PCK of teaching electrochemistry in chemistry teachers: A case in Johannesburg, Gauteng Province, South Africa. *Educacion Quimica*, *25*(3), 336-354.

Sağbilge, K. (2022). *Investigation of science teachers' topic-specific pedagogical content knowledge regarding climate, weather and seasons: from the dual perspective* [Unpublished master's thesis]. Middle East Technical University.

Sandelowski, M. (2010). What's in a name? Qualitative description revisited. *Research in Nursing & Health*, *33*(1), 77-84.

Schmelzing, S., van Driel, J. H., Jüttner, M., Brandenbusch, S., Sandmann, A., & Neuhaus, B. J. (2013). Development, evaluation, and validation of a paper-and-pencil test for measuring two components of biology teachers' pedagogical content knowledge concerning the "cardiovascular system". *International Journal of Science and Mathematics Education*, *11*, 1369-1390.

Schuster, D., Cobern, W. W., Applegate, B., Schwartz, R., Vellom, P., & Undreiu, A. (2007, October). Assessing pedagogical content knowledge of inquiry science teaching. Developing an assessment instrument to support the undergraduate preparation of elementary teachers to teach science as inquiry. *The National STEM Assessment Conference on Assessment of Student Achievement, hosted by the National Science Foundation and Drury University*, Washington DC.

Shulman, L. S. (1986). Those who understand: Knowledge growth in teaching. *Educational Researcher*, *15*, 4-14.

Shulman, L. S. (1987). Knowledge and teaching: Foundations of the new reform. *Harvard Educational Review*, *57*, 1-22.

Simon, M., & Tierney, R. (2011). Use of vignettes in educational research on sensitive teaching functions such as assessment. Retrieved from http://www.icsei.net/icsei2011/Full%20Papers/0153.pdf (accessed February 2015).

Singh, H. (2021). *Understanding and supporting science teachers teaching outside their expertise: Studies of teachers and those who support them* [Unpublished doctoral dissertation]. The University of Georgia.

Sneider, C., Bar, V., & Kavanagh, C. (2011). Learning about seasons: A guide for teachers and curriculum developers. *The Astronomy Education Review*, *10*, 1-22.

Sung, J. Y., & Oh, P. S. (2018). Sixth grade students' content-specific competencies and challenges in learning the seasons through modeling. *Research in Science Education*, *48*, 839-864.

Tsai, C.-C., & Chang, C.-Y. (2005). Lasting effects of instruction guided by the conflict map: Experimental study of learning about the causes of the seasons. *Journal of Research in Science Teaching*, *42*(10), 1089-1111.

Van der Valk, T., & Broekman, H. (1999). The lesson preparation method: A way of investigating pre-service teachers' pedagogical content knowledge. *European Journal of Teacher Education*, *22*, 11-22.

van Driel, J. H., & Abell, S. K. (2010). Science teacher education. In McGraw B., Peterson P. L., Baker E. (Eds.), *International Encyclopedia of Education* (3rd ed., Vol. 7, pp. 712-718). Elsevier.

van Driel, J. H., De Jong, O., & Verloop, N. (2002). The development of pre-service chemistry teachers' pedagogical content knowledge. *Science Education*, *86*, 572-590.

van Driel, J. H., Verloop, N., & De Vos, W. (1998). Developing science teachers' pedagogical content knowledge. *Journal of Research in Science Teaching*, *35*(6), 673-695.

Veal, W. R., & MaKinster, J. G. (1999). Pedagogical content knowledge taxonomies. *Electronic Journal of Science Education*, *3*(4).

Willis, D. G., Sullivan Bolyai, S., Knafl, K., & Cohen, M. Z. (2016). Distinguishing features and similarities between descriptive phenomenological and qualitative description research. *Western Journal of Nursing Research*, *38*(9), 1185-1204.

Appendices

Appendix 1

A Sample Extract of the pPCK Rubric

pPCK		Limited	Basic	Developing	Exemplary
Students' understanding	Students' prior knowledge	-No prior knowledge of the science topic was mentioned. -Several major concepts were not mentioned as prior knowledge about the science topic.	-A few major concepts were mentioned as prior knowledge about the science topic.	-Some major concepts related to the science topic were mentioned sufficiently as prior knowledge.	-As prior knowledge about the science topic, many major concepts were mentioned extensively.
	Students' learning difficulties (misconceptions) and their reasons	-No learning difficulties were mentioned about the science topic. -No misconceptions related to the science topic were identified.	-One or two misconceptions related to the science topic were identified. -The reasons for these misconceptions were not fully explained or partially explained.	-Two or three misconceptions related to the science topic were identified. -The reasons for these identified misconceptions were explained sufficiently.	-All misconceptions related to the science topic were identified. -The reasons for these misconceptions were explained extensively.

Appendix 2

A Sample Extract of the ePCK Rubric

ePCK		Limited	Basic	Developing	Exemplary
Students' understanding	Questioning students' prior knowledge and learning difficulties (misconceptions)	-During the lesson, students' possible preliminary knowledge and misconceptions about the science topic were not questioned.	-During the lesson, students' possible preliminary knowledge and misconceptions about the science topic were partially questioned.	-During the lesson, students' possible preliminary knowledge and misconceptions about the science topic were questioned in an acceptable way.	-During the lesson, students' possible preliminary knowledge and misconceptions about the science topic were adequately questioned.
	Recognizing and acknowledging students' prior knowledge and learning difficulties (misconceptions)	-During the lessons, students' misconceptions related to the science topic were not recognized. -During the lessons, the preliminary knowledge and misconceptions of the students were not taken into consideration.	-During the lesson, a few of the students' misconceptions related to the science topic were recognized. -During the lessons, the preliminary knowledge and misconceptions of the students about the science topic were partially taken into consideration.	-During the lesson, some of the students' misconceptions related to the science topic were recognized. -During the lessons, students' preliminary knowledge and misconceptions related to the science topic were taken into account in an acceptable way.	-During the lesson, many of the students' misconceptions related to the science topic were recognized. -During the lessons, students' preliminary knowledge and misconceptions related to the science topic were adequately taken into consideration.

Journal of Theoretical Educational Science, 17(1), 122-144, January 2024
Kuramsal Eğitimbilim Dergisi, 17(1), 122-144, Ocak 2024
[Online]: http://dergipark.org.tr/akukeg
DOI number: http://doi.org/10.30831/akukeg.1299705

Examining the Development Process of Middle School Students' Knowledge Structures for the Concepts of Melting and Dissolution According to Conceptual Change Theories

Ortaokul Öğrencilerinin Erime ve Çözünme Kavramlarına Yönelik Bilgi Yapılarının Gelişim Sürecinin Kavramsal Değişim Teorilerine Göre İncelenmesi

Mehmet Ali KANDEMİR* **Zeki APAYDIN****

Received: 20 May 2023 **Research Article** **Accepted:** 26 December 2023

ABSTRACT: The aim of this longitudinal research is to examine the development process of scientific and consistent knowledge structures of first, second, and third-grade students in middle school within the concepts of melting and dissolution according to conceptual change theories at the end of the education and training processes carried out in educational institutions. The research is carried out in a district of a metropolitan city in the Marmara Region with the participation of 43 middle school students. Open-ended questions developed for concepts were used to collect data. Friedman test is used to determine whether there is a difference between students' consistent and scientific knowledge structure scores regarding the concepts of melting and dissolution. As there is a significant difference between the students' consistent and scientific knowledge structure scores regarding the concepts of melting and dissolution, the Wilcoxon Signed Rank Test is used for dual comparisons between student levels. In addition, Spearman's correlation coefficient of rank differences is used to determine if there is a relationship between the class levels of the students and the scores of students' consistent and scientific knowledge structure regarding the concepts of melting and dissolution. As a result of the research, it is determined that the students' consistent and scientific knowledge structure scores regarding the concepts of melting and dissolution differ depending on the students' levels. When all the findings obtained from the research are evaluated together, it can be said that the development process of students' knowledge structures regarding the relevant concepts is more compatible with the knowledge in pieces conceptual change theory. Based on the results of the research, some suggestions were made.

Keywords: Knowledge in pieces structure conceptual change theory, theory-like conceptual change theory, mental model.

ÖZ: Bu boylamsal araştırmanın amacı; eğitim kurumlarında gerçekleştirilen eğitim ve öğretim süreçleri sonunda ortaokul birinci, ikinci ve üçüncü sınıf öğrencilerinin erime ve çözünme kavramlarına yönelik bilimsel ve tutarlı bilgi yapılarının gelişim sürecini kavramsal değişim teorilerine göre incelemektir. Araştırma Marmara Bölgesi'ndeki bir büyükşehrin bir ilçesinde 43 ortaokul öğrencisinin katılımı ile gerçekleştirilmiştir. Araştırmada verilerin elde edilmesinde erime ve çözünme kavramlarına yönelik geliştirilen açık uçlu sorular kullanılmıştır. Araştırma verilerinin analizinde Friedman testi, Wilcoxon İşaretli Sıralar Testi ve Spearman'ın sıra farkları korelasyon katsayısı kullanılmıştır. Araştırma sonucunda, öğrencilerin erime ve çözünme kavramlarına ilişkin tutarlı ve bilimsel bilgi yapısı puanlarında, sınıf düzeyine göre farklılıklar olduğu saptanmıştır. Araştırmadan sağlanan tüm bulgular birlikte değerlendirildiğinde öğrencilerin ilgili kavramlara yönelik bilgi yapılarının gelişim sürecinin, daha çok parça nitelikli kavramsal değişim kuramıyla uyumlu olduğu söylenebilir. Araştırma sonuçlarına dayanarak bazı önerilerde bulunulmuştur.

Anahtar kelimeler: Parça nitelikli kavramsal değişim teorisi, teori benzeri kavramsal değişim teorisi, zihinsel model.

* *Corresponding Author*: Primary School Teacher, Ministry of National Education, Balıkesir, Türkiye, mehmetalikandemir10@gmail.com, https://orcid.org/0000-0001-9340-2559
** Prof. Dr., 19 Mayıs University, Samsun, Türkiye, zapaydin@omu.edu.tr, https://orcid.org/0000-0002-6581-4828

Citation Information
Kandemir M. A., & Apaydın, Z. (2024). Examining the development process of middle school students' knowledge structures for the concepts of melting and dissolution according to conceptual change theories. *Kuramsal Eğitimbilim Dergisi [Journal of Theoretical Educational Science], 17*(1), 122-144.

ISSN: 1308-1659

The teaching process of the concepts of melting and dissolution, which are among the concepts included in the science lesson curriculum, is maintained starting from the fourth grade in primary school to the university level (Kirman-Bilgin et al., 2014). Although the students have gone through a long education and training process for the concepts of melting and dissolution, the knowledge structures of the students about the concepts of melting and dissolution may be in conflict with scientific knowledge. In the literature, some findings are reached related to the existence of knowledge structures that are in contrast with scientific and consistent knowledge about the concepts of melting and dissolution in primary school students (Durmus & Donmez-Usta, 2020; Kandemir & Apaydin, 2022, 2023); in university students (Akgun & Aydın, 2009; Kirman-Bilgin et al., 2014; Sen & Yilmaz, 2012) and in middle and high school students as well (Calık & Ayas, 2005). In addition, studies in the international literature reveal that this situation is not unique to our country (Eilks et al., 2007; Goodwin, 2002; Pinto et al., 2023; Ramesh et al., 2020). The contents of the relevant studies in the literature are given below. In their research, Durmus and Donmez-Usta (2020) asked students to draw drawings regarding the concept of melting. It was determined that the drawings made by third-grade primary school students at the scientific level constituted 39.3% of the group, and the drawings made by fourth-grade students at the scientific level constituted 8.1% of the group. In another study in the literature, Kirman-Bilgin et al. (2014) found that science teacher candidates used alternative concepts such as melting, extinction, disappearing, imprisonment, and neutralization instead of the concept of dissolution. Calık and Ayas (2005) examined the knowledge structures of students studying between the seventh and tenth grades regarding selected solution concepts. As a result of the research, it was concluded that some students' answers were inconsistent and that all those depended on the mixture used. In another study, Pinto et al. (2023) examined the knowledge structures of undergraduate and master's degree students regarding the dissolution of oxygen in water. The research revealed that only 11% of the students had scientific knowledge. In their research conducted in primary school, Kandemir and Apaydın (2023) found that, as a result of the pre-test application, the scientific and consistent response score of the students towards the concepts of melting and dissolution was lower than the non-scientific response score.

Not only formal education institutions but also observations of the outside world, communication, and interaction with the social environment are effective in the formation of students' ideas about scientific concepts (Cepni, 2016; Furlough & Gillan, 2018; van Ments & Treur, 2021; Yuzbasıoglu & Kurnaz, 2020). However, the ideas about the concepts obtained in this way are not always correct. Nevertheless, ideas formed as a result of students' misperception of the relevant concept may conflict with scientifically accepted information (Sen & Yilmaz, 2012). Therefore, knowledge structures should be examined in order to know how students construct a concept, to learn their level of knowledge about that concept, and to see whether there are any misconceptions (Kayhan, 2010). The most important reason for analyzing knowledge structures is that knowledge structures are effective in learning new concepts (Ilyas & Saeed, 2018; Sozcu et al., 2016; Vosniadou, 1994). According to the constructivist learning theory, which forms the basis of our science lesson curriculum, meaningful learning occurs by associating the individual's prior knowledge with the new knowledge (Acıkgöz-Ün, 2014; Cepni, 2016). For that reason, it is possible for teachers

to know the knowledge structures of the students about the concept to be learned and to prepare the education-teaching processes by taking these knowledge structures into consideration. Thus, meaningful learning and the quality of education and training process will increase (Ilyas & Saeed, 2018; Malatyali & Yilmaz, 2010; Sozcu et al., 2016). Otherwise, incomplete or incorrect information about concepts may prevent students from learning concepts in a meaningful and permanent way (Ecevit & Simsek, 2017). In this context, knowledge structures are crucial in learning concepts in a meaningful way in the teaching process (Ultay et al., 2017).

Research on conceptual change in science is mainly concerned with two questions: (1) What knowledge of the natural world do pupils have before formal education? (2) How does this knowledge change as children develop and scientific expertise is acquired? After more than two decades of research on the topic, there are still deep disagreements about these questions. Most notable in this context is the question of coherence: Some researchers argue that students' common-sense knowledge of the natural world is coherent. Others argue that it is inconsistent (Sherin et al., 2012). There are different theories in the literature examining students' knowledge structures. However, it is possible to say that research in the literature is mostly grouped into two theories. According to the theory-like conceptual change theory, which is the first of the related theories, students form some naive concepts as a result of their daily experiences. These knowledge structures, which are formed by the individual as a result of their daily experiences, show consistency within themselves (Ozdemir, 2007; Vosniadou, 2012, 2019). However, students' knowledge structure does not have scientific consistency like scientists'. The reason is that students cannot hypothesize and test their ideas (Vosniadou, 2003). These constructs are developed unconsciously as a result of daily experiences and have explanatory power to make consistent predictions and explanations in different domains (Parnafes, 2012). According to the theory mentioned above, students' knowledge structures are tightly interconnected and integrated (Hannust & Kikas, 2007). According to the related theory, when individuals are asked similar questions about a concept, the individuals' answers are based on the same knowledge structures (schema, mental model), so it provides consistency in their answers (Ozdemir, 2007). In addition, according to the theory, students' knowledge structures of the related concept follow a process from a non-scientific consistent structure to a scientific and consistent one (Apaydın, 2020; Ioannides & Vosniadou, 2002; Ozdemir, 2007; Vosniadou, 2003, 2012). According to the theory, students can combine the scientific knowledge that they have encountered in formal education institutions with the naive knowledge that they have created as a result of their daily experiences and can create synthetic structures (Vosniadou, 2012). Even though a synthetic concept may be scientifically incorrect, it still allows the student to advance in the process of acquiring knowledge. A synthetic concept must have both internal consistency and explanatory adequacy (Vosniadou & Skopeliti, 2014). For example, a student comes to a formal education institution knowing that sugar melts in water. At school, he encounters scientific knowledge that sugar dissolves in water. This case creates a mental imbalance for the student. The student combines his prior knowledge of the concept with scientific knowledge in order to get rid of the confusion that surrounds him. As a result, when the student is asked how it is named when water and sugar are

mixed, he is likely to say, "Sugar dissolves in water, and sugar disappears in water." This response proves the existence of the student's synthetic structures.

Synthetic concepts are dynamic, not fixed. It constantly changes as students' information systems evolve. In other words, when the student encounters new information, the student's knowledge structure is reinterpreted. This process continues until complete conceptual change occurs. According to the relevant theory, conceptual change in individuals occurs slowly and gradually. Additionally, in this process, fragmentation and synthetic concepts may occur in students' knowledge structures. The relevant theory emphasizes that rather than telling students that their ideas are wrong and need to be changed, it would be more productive to design instruction that will help them understand that scientific explanations represent a different perspective with more explanatory power compared to their initial understanding (Vosniadou & Skopeliti, 2014).

The other theory that examines students' knowledge structures in the literature is the knowledge in pieces structure conceptual change theory. According to this theory, students have thousands of primitive ideas or understandings that they form as a result of their daily lives, called phenomenological primitives (p-prims). Phenomenological primitives are fragmentary and loosely organized. The fact that students' knowledge structures are fragmented and far from unity may cause them to give inconsistent answers to similar questions about a concept. In addition, according to this theory, students' knowledge structures about a concept follows a process from a fragmented structure to a scientific and consistent structure. The important factor emphasized in this process is the increase in students' knowledge of the relevant concept (Apaydın, 2014, 2020; Clark, 2006; diSessa et al., 2004; Kandemir & Apaydın, 2022; Ozdemir & Clark, 2007; Ozdemir, 2007; Ozdemir, 2018). For example, while a student can state that the salt is dissolved in a salt-water mixture, he can conversely express that sugar melts in a sugar-water mixture. However, dissolution occurs in both mixtures. Phenomenological primitives cannot be removed or destroyed. The loose organization of phenomenological primitives allows a gradual reformation of a student's perception as new information is gathered (Smith et al., 1994). In this theory mentioned before, the student's prior knowledge is valuable because it is emphasized that the advanced knowledge could be reached gradually with the help of rearrangement of the student's prior knowledge. For this reason, they hold the opposite view against the perspective of misconceptions, which is seen as something to be overcome for the development of scientific perception (diSessa, 2014; Smith et al., 1994). According to this theory, conceptual change is a long, gradual, and conceptual process (diSessa, 2014; Furlough & Gillan, 2018; Ozdemir, 2018; Vosniadou, 2019). In the study, the development process of students' scientific and consistent knowledge structures regarding the concepts of melting and dissolution will be discussed within the scope of knowledge in pieces conceptual change theory and theory-like conceptual change theory.

In the literature, there is no longitudinal study that examines the development of scientific and consistent knowledge structures of first, second, and third-grade students in middle school related to the concepts of melting and dissolution. The research is important in terms of filling this gap in the literature and teaching process of the concepts of melting and dissolution. Although there is a long education and training process starting from the 4th grade of primary school and continuing until the

university, student knowledge structures regarding the concepts of melting and dissolution may conflict with scientific knowledge (Kandemir & Apaydın, 2020; Ramesh et al., 2020; Sen & Yılmaz, 2012). The middle school level was preferred in the study because it constitutes the basis for the related concepts after the primary school level; the scope of the related concepts expands at this level, and their teaching is carried out in-depth and in detail. It is a known fact that students' knowledge structures about concepts are very important in the realization of meaningful learning (Kandemir & Apaydın, 2022). Knowledge structures inform teachers about how students construct concepts/concepts (Ozgul et al., 2018). In this direction, teachers also organize the education and training process. Research is important in terms of helping teachers organize the education and training process. In addition, it is observed that qualitative research methods are used in the majority of domestic and foreign studies in the literature, and the existing knowledge structures of students around a concept are examined (Akman, 2013; Apaydın, 2020; Clark, 2006; diSessa et al., 2004; Ozdemir, 2018; Turceotte, 2012; Vosniadou & Brewer, 1992). It is noteworthy that quantitative and longitudinal studies are scarce in the literature (Clark, 2006; Øyehaug & Holt, 2013). It is expected to contribute to the literature about this issue as well. The aim of this longitudinal research is to examine the process of development of scientific and consistent knowledge structures about the concepts of melting and dissolution of first, second, and third-grade students in middle school at the end of the education and training processes carried out in educational institutions. Since the studies on the subject in the literature focus on the knowledge in pieces conceptual change theory and the theory-like conceptual change theory, the development of scientific and coherent knowledge structures for the concepts of melting and dissolution has been examined within the scope of two theories. In this context, the following research questions were determined.

1. Is there a significant relationship between the class levels of the students and the scores obtained from the students' consistent and scientific knowledge structures regarding the concepts of melting and dissolution?

2. Is there a significant difference between the scores obtained from the students' consistent and scientific knowledge structures regarding the concepts of melting and dissolution based on their class levels?

Method

Model of the Research

The research is panel research, which is one type of longitudinal survey research model aiming to examine the development of students' knowledge structures for the concept of melting and dissolution at different class levels. In the related research, measurements are conducted for the same individuals each time at different times. The purpose of the measurements conducted in this way is to examine the changes in the variable during a time process (Buyukozturk et al., 2016; Ozmen & Karamustafaoglu, 2019). In survey research, the existing situation is described and defined as the way it exists in its own conditions, without a trial to change the situation (Karasar, 2015).

Ethical Procedures

All procedures in this study were carried out in accordance with the decision of the ethics committee of 19 Mayıs University Social and Human Sciences dated 29.12.2019 and numbered 2019/455.

Participants

The research was conducted starting from the 2019-2020 academic year up to the 2021-2022 academic year in a district of a metropolitan city in the Marmara Region. The research was carried out with the participation of 43 middle school students. The participants were determined by using the convenience sampling method, which is one of the purposive sampling methods. In the convenience sampling sampling method, the researcher chooses a situation that he or she can easily access (Yıldırım & Şimşek, 2018). In other words, the researcher provides data from the participants that he can easily reach. The method in question is very economical in terms of time, money, and labor. It accelerates the research (Buyukozturk et al., 2016). A two-stage sampling method was used to determine the participants of the study. First, a middle school was selected among the 16 middle schools in the district using the lottery method. Then, two classes were selected from the first grades of the selected middle school by lottery method. The sampling method in both stages of the research was cluster sampling.

Data Collection Tools

In the research, open-ended questions about the concepts of melting and dissolution, which were developed by the researchers in accordance with the grade levels, were used (Appendix-1). Open-ended questions evaluate the development of students' knowledge structures within the scope of theory, like conceptual change theory and knowledge in pieces structure conceptual change theory. While the inconsistent answers given by the students to similar questions were evaluated within the scope of the knowledge in pieces structure conceptual change theory, the consistent answers given by the students to similar questions were evaluated within the scope of the theory-like conceptual change theory. In the process of developing open-ended questions, the literature was first reviewed (Apaydın, 2020; diSessa, 2008; Kandemir & Apaydın, 2022; Ozdemir, 2018; Vosniadou & Brewer, 1992). Then, 10 open-ended items suitable for each grade level were written. Then, the opinions of four field experts, one measurement and evaluation expert, and six science teachers were taken to increase content validity and reliability. Necessary corrections were made accordingly. In order to determine whether the questions were understandable by the students, a pre-pilot application was made to 20 students who would not participate in the research. At the end of the application, feedback was received from the students that the test was comprehensible. The open-ended questions prepared after this stage were applied starting in the 2019-2020 academic year and until the 2021-2022 academic year. Data were obtained by applying similar open-ended questions to the same students regarding the concepts of melting and dissolution at all levels from the first grade to the third grade of middle school. As the grade levels of the students changed, the open-ended questions were rearranged according to the grade level. While preparing open-ended questions at all grade levels, the process of developing open-ended questions above was carried out repeatedly (Atılgan, 2009; Özmen & Karamustafaoglu, 2019).

Data Collection and Analysis

In this longitudinal study, the data was gathered by applying the relevant open-ended questions to the same students three times, first in 2019, in 2020, and then in 2021, from the first grade of middle school to the third grade of middle school. If the students' knowledge structures of concepts are developed from a non-scientific consistent structure to a scientifically consistent structure, they are handled according to the theory-like conceptual change theory. If the students' knowledge structures of concepts show a development from a fragmented (inconsistent) structure to a scientifically consistent structure, they are handled according to the knowledge in pieces structure conceptual change theory. Data analysis of the study was started by entitling the student answer sheets. The open-ended question forms were entitled as 1Ö1, 1Ö2, 1Ö3… 1Ö43 for the students going through their education in the first grade of middle school. The open-ended question forms were entitled 2Ö1, 2Ö2, 2Ö3… 2Ö43 for the students attending the second grade of the secondary school. The open-ended question forms were similarly entitled as 3Ö1, 3Ö2, 3Ö3… 3Ö43 for the students attending the third grade of middle school.

Table 1

Rubric for Students' Responses

Point	Criteria
1	Scientifically consistent answers given by students to similar open-ended questions regarding the same concept were given 1 point. Examples of students' responses are given below.
0	Non-scientific consistent and inconsistent answers were given 0 points. Examples of students' responses are given below.

Scientifically consistent answers given by students to different open-ended questions on the same concept were given one point. Non-scientific consistent and inconsistent answers were given zero points (Table 1). In other words, the students were asked different questions about concepts with the same answers. The consistency or inconsistency of the students' answers to these questions was examined and scored accordingly. Students' scientifically correct and consistent answers to the relevant questions (consistency between the answers given to at least two questions) earned points. Other answers did not gain any points.

Examples of students' responses are given below.

Teacher: What happens to the sugar if water and sugar are mixed?

1Ö3: Sugar dissolves in water.

Teacher: What happens to the salt if water and salt are mixed?

1Ö3: Salt dissolves in water and breaks into small pieces.

In the first question, the student was asked what happens to the sugar in the sugar-water mixture. In the second question, the student was asked what happens to the salt in the salt-water mixture. In response to these questions, the student stated the concept of dissolution. Thus, it was determined that the student answers were scientific

and there was consistency among their answers. One point was given for a scientifically consistent answer.

Teacher: What happens to the sugar if water and sugar are mixed?

1Ö12: Sugar melts in water.

Teacher: What happens to the salt if water and salt are mixed?

1Ö12: Salt melts in water and disappears.

In the first question, the student was asked what happens to the sugar in the sugar-water mixture. In the second question, the student was asked what happens to the salt in the salt-water mixture. The student gave the concept of melting in response to the questions. Thus, it was determined that the student answers were not scientific, and there was consistency among their answers. A non-scientific consistent answer was given zero points.

Teacher: What happens to the sugar if water and sugar are mixed?

1Ö36: Sugar melts in water and disappears.

Teacher: What happens to the salt if water and salt are mixed?

1Ö36: Salt dissolves in water.

In the first question, the student was asked what happens to the sugar in the sugar-water mixture. In the second question, the student was asked what happens to the salt in the salt-water mixture. The student gave the concept of melting as an answer to the first question and the concept of dissolution as an answer to the second question. Thus, it was determined that there was an inconsistency between the student's answers. An inconsistent response was given 0 points.

Then, the data was put in the SPSS 22 data analysis package program. Then, whether the data had a normal distribution or not was examined. According to the results, it was understood that the data did not match with the normality criterion (p<.05). After that, the Friedman test was used to determine whether there was a significant difference between the students' consistent and scientific knowledge structures scores regarding the concepts of melting and dissolution according to their level. As a result of the related test, a significant difference was observed between the scores of the same group. Wilcoxon Signed Ranks Test was used in dual comparisons to determine where the differentiation occurred between the groups. According to the test results obtained, the score ranges used to evaluate the effect of the level variable on the scores are as follows: The effect is low if 0.1<r; it is moderate if 0.3<r; and it is found high if 0.5<r (Cohen, 2013). Also, Spearman's rank correlation coefficient (r_s) is used to determine whether there is a significant relationship between the class level and students' scores of consistent and scientific knowledge structures regarding the concepts of melting and dissolution. Those reasons are effective in the usage of the related method: The data does not have a normal distribution, and one of the variables is at the level of the ranking scale while the other one is at the level of the equally spaced scale. The correlation coefficient ranges are as follows: high correlation is between 1-0.70; moderate level is between 0.69-0.30; and low level can be attributed as between 0.29-0.00. The determination coefficient is calculated by squaring the correlation coefficient found (R^2). Determination correlation helps to figure out how much a change in one of the variables is explained by the other variable (Kilmen, 2015).

Results

The research findings were examined within the scope of the knowledge in pieces conceptual change theory and theory-like conceptual change in the development of scientific and consistent knowledge structures of students for the concepts of melting and dissolution.

Within the scope of the first research question, the existence or absence of a significant relationship between the student class levels and the scores obtained from the students' consistent and scientific knowledge structures regarding the concepts of melting and dissolution was investigated. Obtained findings are given below.

Table 2

Spearman's Rank Differences Correlation Coefficient Test Results

		Points
Class level	Correlation Coefficient	.305
	Sig. (2-tailed)	.000*
	N	129

*p<.05

According to the test result in Table 2, there is a moderate, positive, and significant relationship between the class level of the students and the scores of scientific knowledge structures (r_s=.305, p<.05). According to this result, it can be stated that as the student level increases, the students' consistent and scientific knowledge scores about the concepts of melting and dissolution increase. The determination coefficient is calculated by squaring the correlation coefficient obtained. R^2=.305*.305=.09. Based on this result, it can be stated that approximately 9% of the variability in the consistent and scientific knowledge structure scores can be explained by the student class level variable.

Within the scope of the second research question, the scores obtained from the consistent and scientific knowledge structures by the students related to the concept of melting and dissolution were compared at the first, second, and third class levels of middle school. The findings obtained in this term are given below:

Table 3

Friedman Test (X^2) Results Between Different Class Levels

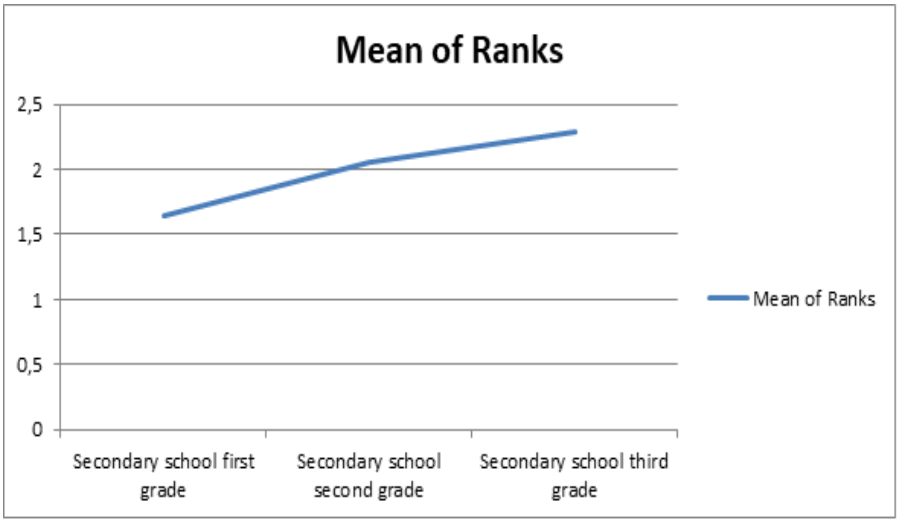
Class levels	*N*	Mean of Ranks	*sd*	x^2	*p*
Middle school first-grade	43	1.65	2	19.872	.000*
Middle school second-grade	43	2.06			
Middle school third-grade	43	2.29			

*p<0.05

According to the results of the analysis in Table 3, a significant difference is found between the scores obtained from the students' scientific and consistent knowledge structures for the concepts of melting and dissolution in the first grade of middle school, second grade of the middle school, and third grade of the middle school (X^2=19.872, p<.05). In order to find out between which grades this difference occurs, dual comparisons are made between middle school first grade-middle, school second grade, middle school first grade-middle school third grade, middle school third grade-middle school second grade. Wilcoxon Signed Rank Test is used for dual comparisons.

Figure 1

Line Graph of the Mean of Ranks for Different Class Levels



In Figure 1, it is observed that the mean rank of the middle school first grade is 1.65, the mean rank of the middle school is 2.06, and the mean rank of the middle school third grade is 2.29. Based on these findings, it can be stated that students' scientific and consistent knowledge structures about the concept of melting and dissolution make progress slowly and gradually. In addition to that, as the class level increases, it is observed that scientific and consistent knowledge structures about the concepts of melting and dissolution increase, too.

Table 4

The Test Result of the Comparison of the Middle School First Class and Second Grade Scores

Class levels		*N*	Mean of Rank	Sum of Ranks	*Z*	*p*	*r*
Middle school first-grade Middle school second-grade	Negative Ranks	1[a]	5.50	5.50	-3.087	.002	.46
	Positive Ranks	13[b]	7.65	99.50			
	Ties	29[c]					
	Total	43					

p<.05

a. Middle school second-grade<Middle school first-grade

b. Middle school second-grade>Middle school first-grade

c. Middle school second-grade=Middle school first-grade

When the test analysis results in Table 4 are examined, it is seen that 13 results are positive, 29 results are equal, and one result is negative if the consistent and scientific knowledge structures scores of the second-grade students in middle school are subtracted from the consistent and scientific knowledge structure scores of the first-grade students in middle school. Based on these findings, it can be stated that the scores of the second-grade (middle school) students' consistent and scientific knowledge structures are higher than the knowledge structure scores of the first-grade students. In addition, according to the Wilcoxon Signed Rank Test result, a significant difference is found in the dual comparison between the consistent and scientific knowledge structure scores of the second-grade students in middle school and the consistent and scientific knowledge structures scores of the first-grade students in middle school (Z=-3.087, p<.05 r=.46). In order to determine which class level is in favor of this difference, the mean values of the rows are checked. A higher mean of ranks means higher scores. While the mean row value for the measurements of the first graders in middle school is = 5.50, the average row value for the measurements of the second graders in middle school is = 7.65. Therefore, this difference is in favor of the consistent and scientific knowledge structure scores of the second-grade students in middle school. In other words, the student class level variable has a significant effect on students' consistent and scientific knowledge structures. This effect is moderate (r=.46).

Table 5

The Test Result of the Comparison of Middle School Third Class-Middle School First Grade Scores

Class levels		*N*	Mean of Rank	Sum of Ranks	*Z*	*p*	*r*
Middle school first-grade	Negative Ranks	2[d]	7.50	15.00	-3.740	.000	.56
	Positive Ranks	20[e]	11.90	238.00			
Middle school third-grade	Ties	21[f]					
	Total	43					

p<.05

d. Middle school third-grade<Middle school first-grade

e. Middle school third-grade>Middle school first-grade

f. Middle school third-grade=Middle school first-grade

When the test analysis result in Table 5 is examined, it is seen that 20 results are positive, 21 results are equal, and two results are negative as long as the consistent and scientific knowledge structures scores of the first-grade students in middle school are subtracted from the coherent and scientific knowledge structures scores of the third-grade students in middle school. Based on this result, it can be stated that the consistent and scientific knowledge structures scores of the third-grade students in middle school are higher than the consistent and scientific knowledge structures scores of the first

graders in the middle. Furthermore, according to the Wilcoxon Signed Rank Test result, a significant difference is found between the consistent and scientific knowledge structures scores of the third-grade students in middle school and the consistent and scientific knowledge structures scores of the first graders in middle school (*Z*=-3.740, *p*<.05, *r*=.56). In order to specify which class level is in favor of this difference, the mean values of the rows are checked. A higher mean of ranks means higher scores. When the mean of ranks for the measurements of the first graders in middle school is = 7.50, the mean of ranks for the third graders in middle school is = 11.90. Thus, this difference is in favor of coherent and scientific knowledge structure scores of the third-grade students in middle school. That is to say, the student-level variable has a significant effect on students' consistent and scientific knowledge structures. This effect is high (*r*=.56).

Table 6

The Test Result of the Comparison of Middle School Third Class and Second Grade Scores

Class levels		*N*	Mean of Rank	Sum of Ranks	*Z*	*p*	*r*
Middle school second-grade Middle school third-grade	Negative Ranks	5[g]	6.50	32.50	-2.173	.030	.32
	Positive Ranks	12[h]	10.04	120.50			
	Ties	26[i]					
	Total	43					

p<.05

g. Middle school third-grade <Middle school second-grade

h. Middle school third-grade>Middle school second-grade

i. Middle school third-grade=Middle school second-grade

When the test analysis result in the table is examined, it is seen that 12 results are positive, 26 results are equal, and five results are negative if the consistent and scientific knowledge structure scores of the second-grade students in middle school are subtracted from the coherent and scientific knowledge structures scores of the third-grade students in middle school. Based on this result, it is clear that the consistent and scientific knowledge structure scores of the third graders in middle school are higher than those of the second graders in middle school. In addition, according to the Wilcoxon Signed Rank Test result, a significant difference is observed between the consistent and scientific knowledge structure scores of the third-grade students in middle school and the consistent and scientific knowledge structure scores of the second-grade students in middle school (*Z*=-2.173, *p*<.05, *r*=.32). In order to determine which class level is in favor of this difference, the mean values of the rows are checked. A higher mean of ranks means higher scores. When the mean of ranks for the measurements of the second graders in middle school is = 6.50, the mean of ranks for the measurements of the third graders in middle school is = 10.04. Therefore, this difference is in favor of coherent and scientific knowledge structure scores of the third-grade students in middle school. In other words, the student-level variable has a

significant effect on students' consistent and scientific knowledge structures. This effect is moderate (r=.32).

Table 7

The Number of Students Giving Consistent and Inconsistent Answers According to Grade Level Regarding the Related Concepts

	Middle school first-grade	Middle school second-grade	Middle school third-grade
Number of students who responded scientifically consistent	5	18	30
Number of students with non-scientific consistent response	13	10	9
Number of students with inconsistent response	25	15	4

When Table 7 is analysed, it is observed that in the first grade of middle school, the number of students giving scientifically consistent answers was five, the number of students giving non-scientifically consistent answers was 13, and the number of students giving inconsistent answers was 25. In the second grade of middle school, the number of students giving scientifically consistent answers was 18, the number of students giving non-scientifically consistent answers was 10, and the number of students giving inconsistent answers was 15. In the third grade of middle school, the number of students giving scientifically consistent answers was 30, the number of students giving non-scientifically consistent answers was nine, and the number of students giving inconsistent answers was four.

Discussion and Conclusion

In this part of the study, the findings obtained from the research questions are compared with a few research findings in the literature. In the research, the findings related to the scientific and consistent knowledge structure and the process of development of the students regarding the related concepts were discussed within the scope of the knowledge in pieces conceptual change theory and the theory-like conceptual change theory.

Within the scope of the first research question, whether there is a significant relationship between the students' class levels and the scientific and consistent knowledge structure scores related to the concept of melting and dissolution is investigated. According to the findings in Table 2, there is a moderate, positive, and significant relationship between student level and consistent and scientific knowledge structure scores (r_s=.305, p<.05). According to this result, it can be said that as the students' level increases, the scores of consistent and scientific knowledge structures related to the concepts of melting and dissolution increase (Kilmen, 2015). In addition, Table 7 shows that the number of students giving scientific and coherent answers

increased as the grade level increased, which supports this finding. This research finding indicates an increase in students' scientific knowledge of the related concept, which is one of the important factors in the formation of scientific and consistent knowledge structures about a concept, and also stated in the related conceptual change theories. Therefore, these findings are consistent with conceptual change theories (diSessa, 2014; Vosniadou, 2012).

The scores obtained in the first, second, and third grades of middle school regarding the scientific and consistent knowledge structures of the students are compared within the scope of the second research question. According to the analysis result in Table 3, a significant difference is found between the scores obtained from the students' consistent and scientific knowledge structures related to the concepts of dissolution and melting in the first, second, and third grades of middle school (X^2=19.872, p<.05). Wilcoxon Signed Rank Test was used for dual comparisons between the first and the second grade of middle school, the first and third grade of middle school, the first and third grades of middle school, and the third and second grades of the middle school in order to find out between which grades this difference is.

According to the first comparison results (Table 4), a significant difference is found between the consistent and scientific knowledge structure scores of the second graders of the middle school and the consistent and scientific knowledge structure scores of the first-grade students of the middle school (Z=-3.087, p<.05, r=.46). This difference is in favor of the consistent and scientific knowledge structure scores of the second-grade students of the middle school with a high mean of ranks (Table 3). The student-level variable has a significant effect on students' consistent and scientific knowledge structures. This effect is moderate (r=.46).

For the second comparison (Table 5), a significant difference is found between the consistent and scientific knowledge structure scores of the third grade of the middle school and the consistent and scientific knowledge structures scores of the first grade of the middle school according to the test results (Z=-3.740, p<.05, r=.56). This difference is in favor of the consistent and scientific knowledge structures scores of the third-grade students of the middle school with a high mean of ranks (Table 4). The student class level variable has a significant effect on students' consistent and scientific knowledge structures. This effect is high (r=.56).

In the third comparison (Table 6), based on the test results, a significant difference is found between the consistent and scientific knowledge structure scores of the third-grade students of the middle school and the consistent and scientific knowledge structures scores of the second-grade students of the middle school (Z=-2.173, p<.05, r=.32). This difference is in favor of the consistent and scientific knowledge structure scores of the third-grade students of the middle school with a high mean of ranks (Table 6). The student class level variable has a significant effect on students' consistent and scientific knowledge structures. This effect is moderate (r=.32).

When the findings obtained in the research and the graph in Figure 1 are evaluated together, it can be stated that the students' consistent and scientific knowledge structures regarding the concept of melting and dissolution become more scientific and consistent as the students' class level increases. The findings of the research are in harmony with the statement, "Students' knowledge structures about a concept become scientific and consistent as their knowledge about that concept increases," which is

emphasized in the knowledge in pieces conceptual change theory and the theory-like conceptual change theory (diSessa, 2014; Vosniadou, 2012).

Moreover, when Table 4, Table 5, and Table 6 are examined, the closeness of the mean of ranks indicates that the scientific and consistent process of students' knowledge of the concepts of melting and dissolution takes place slowly and gradually. This research finding is in harmony with the related conceptual change theories. Also in the literature, Clark (2006), diSessa et al. (2004), Ozdemir (2007), Ozdemir (2018), diSessa (2014), Smith et al. (1994), and Vosniadou (2012, 2019) state in their research that scientific and consistent knowledge structures in students are formed as a result of a slow, long and gradual process. These statements support the research findings. However, when Sagdıc and Sahin's (2023) research on the phases of the moon is compared to our study, it is observed that students make conceptual changes in a shorter time. This contradiction can be attributed to the fact that many factors are effective in conceptual change.

When Table 7 is analysed, it is seen that the number of students giving inconsistent answers has the highest frequency among the first-grade middle school students. It is also noteworthy that the number of students giving inconsistent answers decreases as the grade level increases. In the related table, it is seen that the number of students giving scientific and consistent answers increased, while the number of students giving non-scientific and consistent answers did not change much. These findings show that students' knowledge structures about the related concepts transformed from a fragmented structure to a coherent scientific structure. When all the findings of the study are evaluated together, it can be said that the development process of students' knowledge structures for the related concepts is more compatible with the knowledge process in the theory of the knowledge in pieces conceptual change theory.

In the longitudinal studies on the subject in the literature, Øyehaug and Holt (2013) conducted two years of longitudinal research and found that students' knowledge structures regarding the concepts of the substance and chemical reaction are fragmented and inconsistent; it gradually develops towards an integrated and coherent structure. Another research in the literature is Clark's (2006) longitudinal study on students' knowledge structures on the subject of thermodynamics. It shows that students' knowledge structures about thermal equilibrium evolve from an inconsistent structure to an integrated and consistent structure. The common finding of the studies given above in the literature is that students' knowledge structures become more scientific and consistent over time. With the related finding, it can be stated that the mean ranks of the students' scientific and coherent knowledge structure scores in the related table (Table 2) obtained from the research show parallelism to the finding obtained from the first grade of the middle school to the third grade of it.

Although there are few longitudinal studies within the scope of this subject in the literature, there are some studies whose findings are compatible with our research at different class levels. In the literature, Ozturk and Doganay (2013), in their research, examined the students' knowledge structures about the Earth's shape and gravity from the first grade to the fourth grade of the middle school. As a result of the research, they came up with the result that there is an increase in the scientific and consistent knowledge structures of the students from the first grade to the fourth grade of the middle school. Bilir and Karacam (2021), who examined the knowledge structures of

pre-service science teachers regarding the concept of chemical reaction, revealed that the knowledge structures of the students become more scientific and consistent as the education level increases. Durmus and Donmez-Usta's (2020) research makes it clear that the knowledge structures of the concept of melting are more scientific for fourth-grade students compared to third-grade students. The study by Atasoy et al. (2020), which includes all levels of primary school consisting of one hundred students for each grade, demonstrates that their students' knowledge of marine environments is parallel to their education level. The common result of the studies above in the literature supports the finding of this research: "As the education level of the students increases, the knowledge structures of the students become more scientific and consistent." The relevant finding coincides with the characteristics of the scientific and consistent knowledge development process of students stated by the relevant conceptual change theories. Another study in the literature by Ulu and Ocak (2018) determined that the knowledge structures of the fourth grade of primary school and the first grade of middle school science lessons did not differ. This finding contradicts the research finding. This contradiction can be attributed to the slow and gradual realization of the knowledge structure development process when the mean ranks of the students are examined, as found in our research findings (Table 3).

As the student class level increases, many factors can be effective in increasing the scientific and consistent knowledge structures of the students. Relevant factors can be expressed as follows: The field and pedagogical competence of the teacher who guides the education and training process (Fulmer, 2013), the models, methods, and techniques used in the education and training process (Sozcu & Aydinozu, 2018; Turk at al., 2016), developing students' theoretical reasoning skills as students' cognitive development shifts from concrete operations to abstract operations (Cepni, 2016; Lawson, 1995), the increase in students' knowledge about the concept and their accessibility to scientific knowledge more (diSessa, 2014; Ozdemir, 2007; Ozdemir, 2018) and students' reaching to sufficient or insufficient number of examples regarding the concept (Saygili, 2015; Tas & Karatas, 2012).

It should be noted that the findings from the research are valid for the concepts of melting dissolution and the participants taking part in the study. In another study, different findings may be concluded about the melting dissolution or other concepts in the science curriculum. As stated above, many factors are effective in increasing the scientific and consistent knowledge structures of students about a concept. As mentioned above, although this research is a longitudinal study, it does not show that students' conceptual change process for all concepts will be realized in this way. However, all research on students' knowledge structures contributes to a better understanding of students' knowledge structures, and teachers, who are the guides of the teaching process, can design the teaching process in accordance with this process (Sherin et al., 2012). As a natural consequence of this situation, it becomes easier for students to learn concepts in a meaningful way and to acquire scientific and coherent knowledge structures.

In summary, the findings of the research are found parallel to the statement, "Students' knowledge structures about a concept gradually and slowly become scientific and consistent as their knowledge about that concept increases," which is emphasized in the knowledge in pieces conceptual change theory and the theory-like conceptual

change theory. When all the research findings are evaluated together, it can be said that the development process of the students' knowledge structures for the related concepts is more in line with the knowledge in pieces conceptual change theory.

Suggestions

Based on the results obtained, conceptual change theories can be taught to prospective teachers in education faculties as a course at the undergraduate level. By signing a protocol between the Ministry of National Education and the Higher Education Board, in-service training on the theories of conceptual change can be given to teachers. Science curricula and textbooks can be arranged according to conceptual change theories. Similar longitudinal studies could be carried out by increasing the number of participants and length of the research period for other concepts in the science curriculum.

Acknowledgements

The research was not funded by any institution.

Statement of Responsibility

Mehmet Ali Kandemir: Literature review, data collection tool development, data collection and analysis, expert opinion, writing the text, visualization, conclusion and discussion.

Zeki Apaydın: Literature review, text review and editing, data collection tool development, data collection and analysis, expert opinion, conclusion and discussion.

Conflicts of Interest

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Author Bios:

The corresponding author, Mehmet Ali Kandemir, is a primary school teacher doctoral student. He works as a classroom teacher in the Ministry of National Education. He has 18 articles on argumentation, conceptual change theories, active learning, inclusive education.

Second author Prof. Dr. Zeki Apaydın is a lecturer at Ondokuz Mayis University, Faculty of Education, Department of Basic Education. He has more than 30 articles on argumentation, modelling, conceptual change theories, active learning, inclusive education.

References

Acıkgöz-Ün, K. (2014). *Aktif öğrenme*. Biliş Yayınevi.

Akgun, A., & Aydın, M. (2009). Erime ve çözünme konusundaki kavram yanılgılarının ve bilgi eksikliklerinin giderilmesinde yapılandırmacı öğrenme yaklaşımına dayalı grup çalışmalarının kullanılması. *Elektronik Sosyal Bilimler Dergisi, 8*(27), 190-201.

Akman, E. (2013). *İlköğretim öğrencilerinin ışık kavramına yönelik bilgi yapılarının kavramsal değişim teorilerine göre analizi* [Yüksek lisans tezi]. Ondokuz Mayıs Üniversitesi.

Apaydın, Z. (2014). Ortaokul öğrencilerinin suyun kaldırma kuvveti kavramına yönelik bilgi yapıları: Görüngübilimsel bir ilksel olarak yüzme. *Eğitim ve Bilim, 39*(174), 402-424. http://dx.doi.org/10.15390/EB.2014.3258

Apaydın, Z. (2020). A phenomenological study in the context of conceptual change theories about buoyancy. *International Journal of Education Technology and Scientific Researches, 5*(13), 1711-1789. http://dx.doi.org/10.35826/ijetsar.241

Atasoy, V., Ahi, B., & Balci, S. (2020). What do primary school students' drawings tell us about their mental models on marine environments? *International Journal of Science Education, 42*(17), 2959-2979. https://doi.org/10.1080/09500693.2020.1846821

Atılgan, H. (Ed.). (2009). *Eğitimde ve ölçme değerlendirme*. Anı Yayıncılık.

Bilir, V., & Karacam, S. (2021). Evaluation of mental models of prospective science teachers on chemical reactions. *Journal of Pedagogical Research, 5*(1), 258-274.

Buyukozturk, S., Cakmak, E. K., Akgun, Ö. E., Karadeniz, S., & Demirel, F. (2016). *Bilimsel araştırma yöntemleri*. Pegem Akademi.

Calık, M., & Ayas, A. (2005). 7.-10. sınıf öğrencilerinin seçilen çözelti kavramlarıyla ilgili anlamalarının farklı karışımlar üzerinde incelenmesi. *Türk Eğitim Bilimleri Dergisi, 3*(3), 329-349.

Cepni, S. (Ed.). (2016). *Fen ve teknoloji öğretimi*. Pegem Akademi.

Clark, D. B. (2006). Longitudinal conceptual change in students' understanding of thermal equilibrium: An examination of the process of conceptual restructuring. *Cognition Instruct., 24*(4), 467-563. https://doi.org/10.1207/s1532690xci2404_3

Cohen, J. (2013*). Statistical power analysis for the behavioral sciences*. Academic Press.

diSessa, A. A. (2014). A history of conceptual change research: Threads and fault lines. *In The Cambridge Handbook of the Learning Sciences (Second ed.)* (pp. 88-108). Cambridge University Press. http://dx.doi.org/10.1017/CBO9781139519526.007

diSessa, A. A., Gillespie, N., & Esterly, J. (2004). Coherence versus fragmentation in the development of the concept of force. *Cognitive Science, 28*(6), 843-900. https://doi.org/10.1207/s15516709cog2806_1

Durmus, T., & Donmez-Usta, N. (2020). İlköğretim öğrencilerinin erime kavramı ile ilgili zihinsel modelleri. *International Social Mentality and Researcher Thinkers Journal, 6*(38), 2216-2231.

Ecevit, T., & Simsek, Ö. P. (2017). Öğretmenlerin fen kavram öğretimleri, kavram yanılgılarını saptama ve giderme çalışmalarının değerlendirilmesi. *Elementary Education Online, 16*(1), 129-150.

Eilks, I., Moellering, J., & Valanides, N. (2007). Seventh-grade students' understanding of chemical reactions: reflections from an action research ınterview study. *Eurasia Journal of Mathematics, Science & Technology Education, 3*(4), 271-286. https://doi.org/10.12973/ejmste/75408

Fulmer, G. W. (2013). Constraints on conceptual change: how elementary teachers' attitudes and understanding of conceptual change relate to changes in students' conceptions. *Journal of Science Teacher Education, 24*(7), 1219-1236. https://doi.org/10.1007/s10972-013-9334-3

Furlough, C. S., & Gillan, D. J. (2018). Mental models: Structural differences and the role of experience. *Journal of Cognitive Engineering and Decision Making*, *12*(4), 269-287. https://doi.org/10.1177/1555343418773236

Goodwin, A. (2002). Is salt melting when it dissolves in water. *Journal of Chemical Education, 9*(3), 93-96.

Hannust, T., & Kikas, E. (2007). Children's knowledge of astronomy and its change in the course of learning. *Early Childhood Research Quarterly*, *22*(1), 89-104. http://doi.org/10.1016/j.ecresq.2006.11.001

Ilyas, A., & Saeed, M. (2018). Exploring teachers understanding about misconceptions of secondary grade chemistry students. *International Journal for Cross-Disciplinary Subjects in Education (IJCDSE)*, *9*(1), 3323-3328.

Ioannides, C., & Vosniadou, S. (2002). The changing of force. *Cognitive Science Quarterly*, *2*(1), 5-61.

Kandemir, M. A., & Apaydın, Z. (2020). Sınıf öğretmenlerinin fen bilimleri dersinde öğrencilerin sahip olduğu kavram yanılgılarını belirlemelerine ve gidermelerine yönelik bir değerlendirme. *Türkiye Bilimsel Araştırmalar Dergisi, 5*(2), 82-97.

Kandemir, M. A., & Apaydın, Z. (2022). Analysis of primary school students' knowledge structures regarding the movements of the earth according to conceptual change theories. *Cukurova University Faculty of Education Journal, 51*(2), 825-869. https://doi.org/10.14812/cuefd.934087

Kandemir, M. A., & Apaydın, Z. (2023). Jigsaw II tekniğinin dördüncü sınıf öğrencilerinin erime ve çözünme kavramlarına yönelik bilgi yapılarına etkisi. *Abant İzzet Baysal Üniversitesi Eğitim Fakültesi Dergisi, 23*(2), 585-603. http://doi.org/10.17240/aibuefd.2023..-1102529

Karasar, N. (2015). *Bilimsel araştırma yöntemi*. Nobel Yayıncılık.

Kayhan, C. H. (2010). Model ve zihinsel modeller. *Erzincan Üniversitesi Sosyal Bilimler Enstitüsü Dergisi*, *2*(3), 407-422.

Kilmen, S. (2015). *Eğitim araştırmaları için SPSS uygulamalı istatistik*. Edge Akademi.

Kirman-Bilgin, A., Nas, S. E., & Akbulut, H. İ. (2014). Öğretmen adaylarının "çözünürlük" konusuna yönelik alternatif kavramlarının belirlenmesi. *Uludağ Üniversitesi Eğitim Fakültesi Dergisi, 27*(2), 371-392.

Lawson, A. E. (1995). *Science teaching and the development of thinking*. Wadsworth.

Malatyali, E., & Yilmaz, K. (2010). Yapılandırmacı öğrenme sürecinde kavramlar ve önemi: Kavramların pedagojik açıdan incelenmesi. *Uluslararası Sosyal Araştırmalar Dergisi*, *3*(14), 320-332.

Ozdemir, G. (2007). Öğrencilerin kuvvet kavramına ilişkin bilgi yapılarının bir analizi. *Mehmet Akif Ersoy Üniversitesi Eğitim Fakültesi Dergisi, 8*(14), 37-54.

Ozdemir, G., & Clark, D. B. (2007). An overview of conceptual change theories. *Eurasia Journal of Mathematics, Science ve Technology Education, 3*(4), 351-361. https://doi.org/10.12973/ejmste/75414

Ozdemir, M. (2018). *Ortaokul öğrencilerinin buharlaşma ve kaynama kavramlarına ilişkin bilgi yapılarının analizi* [Yüksek lisans tezi]. Niğde Ömer Halisdemir Üniversitesi.

Ozgul, F., Kangalgil, M., Diker, G., & Yamen, E. (2018). Evaluation of the constructivist learning environments of physical education teacher candidates. *European Journal of Educational Research, 7*(3), 653-658.

Ozmen, H., & Karamustafaoglu, O. (2019). *Eğitimde araştırma yöntemleri.* Pegem Akademi.

Ozturk, A., & Doganay, A. (2013). Primary school 5th and 8th graders' understanding and mental models about the shape of the world and gravity. *Educational Sciences: Theory and Practice, 13*(4), 2469-2476.

Øyehaug, A. B., & Holt, A. (2013). Students' understanding of the nature of matter and chemical reactions-a longitudinal study of conceptual restructuring. *Chemistry Education Research and Practice, 14*(4), 450-467. https://doi.org/10.1039/C3RP00027C

Parnafes, O. (2012). Developing explanations and developing understanding: Students explain the phases of the moon using visual representations. *Cognition and Instruction, 30*(4), 359-403. https://doi.org/10.1080/07370008.2012.716885

Pinto, G., Castro-Acuña, C. M., López-Hernández, I., & Alcázar Montero, V. (2023). Learning difficulties in the interpretation of matter at the molecular level by university students—A case study: Dissolution of oxygen in water. *Education Sciences, 13*(8), 781. https://doi.org/10.3390/educsci13080781

Ramesh, M., Victor, S. R., & Nagaraju, M. T. V. (2020). Misconceptions in certain science concepts among tribal students. *An International Bilingual Peer Reviewed Refereed Research Journal, 10*(40), 24-28.

Sagdıc, A., & Sahin, E. (2023). Examination of dynamic mental constructs and their change regarding phases of the moon. *Journal of Science Learning, 6*(2), 181-193. https://doi.org/10.17509/jsl.v6i2.51686

Saygili, G. (2015). *İlkokulda kullanılan strateji, yöntem ve teknikler*. Pegem Akademi.

Sen, S., & Yilmaz, A. (2012). Erime ve çözünmeyle ilgili kavram yanılgılarının ontoloji temelinde İncelenmesi. *Amasya Üniversitesi Eğitim Fakültesi Dergisi, 1*(1), 54-72.

Sherin, B. L., Krakowski, M., & Lee, V. R. (2012). Some assembly required: How scientific explanations are constructed during clinical interviews. *Journal of Research in Science Teaching, 49*(2), 166-198. https://doi.org/10.1002/tea.20455

Smith, J. P., diSessa, A. A., & Roschelle, J. (1994). Misconceptions reconceived: A Constructivist analysis of knowledge in transition. *Journal of the Learning Sciences, 3*(2), 115-163. https://doi.org/10.1207/s15327809jls0302_1

Sozcu, U., & Aydinozu, D. (2018). 7. sınıf öğrencilerinin bilimsellik değerine ilişkin zihinsel modellerindeki değişim. *Kastamonu Education Journal, 26*(2), 589-597. https://doi.org/10.24106/kefdergi.389882

Sozcu, U., Kildan, A. O., Aydinozu, D., & Ibret, B. Ü. (2016). Bilimsellik değerine ilişkin zihinsel modellerin değişiminin çeşitli değişkenler açısından incelenmesi. *Cumhuriyet International Journal of Education, 5*(2), 9-22.

Tas, M. A., & Karatas, M. K. (2012). Öğretim hedefleri ünitesindeki bilişsel alanda davranışsal amaç kavramının öğretiminde sunulan örneklerin çeşitliliğinin değerlendirilmesi. *Türk Eğitim Bilimleri Dergisi*, *10*(3), 541-583.

Turceotte, S. (2012). Computer-supported collobrative inquiry on buoyancy: A discourse analysis supporting the "pieces" position on conceptual change. *Journal of Science Education, 21*, 808-825.

Turk, C., Kalkan, H., Kıroglu, K., & Ocak-Iskeleli, N. (2016). Elementary school students' mental models about formation of seasons: a cross sectional study. *Journal of Education and Learning, 5*(1), 7-30.

Ultay, E., Donmez Usta, N., & Durmus, T. (2017). Eğitim alanında yapılan zihinsel model çalışmalarının betimsel içerik analizi. *Yaşadıkça Eğitim, 31*(1), 21-40.

Ulu, H., & Ocak, I. (2018). İlköğretim öğrencilerinin fen öğretimine yönelik zihinsel modellerinin incelenmesi. *Electronic Turkish Studies, 13*(11), 1367-1388.

van Ments, L., & Treur, J. (2021). Reflections on dynamics, adaptation and control: A cognitive architecture for mental models. *Cognitive Systems Research, 70*, 1-9. https://doi.org/10.1016/j.cogsys.2021.06.004

Vosniadou, S. (1994). Capturing and modeling the process of conceptual change. *Learning and Instruction, 4*(1), 45-69. https://doi.org/10.1016/0959-4752(94)90018-3

Vosniadou, S. (2003). Exploring the relationships between conceptual change and intentional learning. In G. M. Sinatra & P. R. Pintrich (Eds.), *Intentional conceptual change* (pp. 377-406). Lawrence Erlbaum Associates.

Vosniadou, S. (2012). Reframing the classical approach to conceptual change: preconceptions, misconceptions and synthetic models. Fraser B., Tobin K., & McRobbie C. (Eds), *Second International Handbook of Science Education. Springer International Handbooks of Education* (pp. 119-130). Springer. https://doi.org/10.1007/978-1-4020-9041-7_10

Vosniadou, S. (2019). The development of students' understanding of science. *Frontiers in Education, 4*(32), 1-6. https://doi.org/10.3389/feduc.2019.00032

Vosniadou, S., & Brewer, W. F. (1992). Mental models of the earth: A study of conceptual change in childhood. *Cognitive Psychology, 24*(4), 535-585. https://doi.org/10.1016/0010-0285(92)90018-W

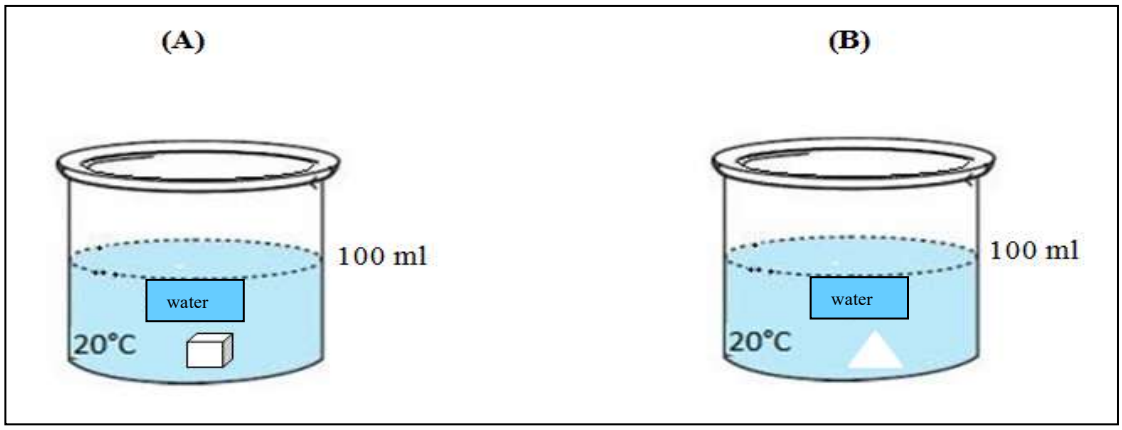
Vosniadou, S., & Skopeliti, I. (2014). Conceptual change from the framework theory side of the fence. *Science & Education, 23*, 1427-1445. https://doi.org/10.1007/s11191-013-9640-3

Yıldırım, A., & Şimşek, H. (2016). *Sosyal bilimlerde nitel araştırma yöntemleri*. Seçkin Yayıncılık.

Yuzbasıoglu, M. K., & Kurnaz, M. A. (2020). Ses hakkında öğrenci zihinsel modellerinin belirlenmesi. *Anadolu Üniversitesi Eğitim Fakültesi Dergisi (AUJEF), 4*(3), 254-27. https://doi.org/10.34056/aujef.687236

Appendix-1

Open-Ended Questionnaire for the Concepts of Melting and Dissolution (Question Examples)



Note: Conditions are equal.

1. (A) What happens to the sugar if we put a sugar cube in the graduated cylinder and mix it?

……

……

2. (B) What happens to the salt if we put salt in the graduated cylinder and mix it?

……

……

3. (A) Where do you think the sugar mixed in the graduated cylinder is? How do you prove the presence or absence of sugar?

……

……

4. (A) Did the sugar maintain its solid condition? Why?

……

5. (B) Where do you think the salt mixed in the graduated cylinder is? How do you prove the presence or absence of salt?

……

……

6. (B) Did the salt retain its solid state? Why?

……

……

7. What can be said about the mass of the sugar if we put a sugar cube into a graduated cylinder with water and mix it?

……

……

8. What we put a sugar cube into a graduated cylinder with water in it and mix it, what can be said about the volume of the sugar?

……

……

9. There are two differently graduated cylinders with the same amount of water and the same temperature. First, the same amount of sugar cubes were placed into these graduated cylinders. Then, some water with the same temperature was added into one of the graduated cylinders. Accordingly, in which graduated cylinder does the sugar dissolve faster?

……

……

10. If we put the same amounts of rock salt and powdered salt into two different graduated cylinders containing the same amount of water at the same temperature, in which graduated cylinder will the salt dissolve faster?

……

……

Journal of Theoretical Educational Science, 17(1), 145-168, January 2024
Kuramsal Eğitimbilim Dergisi, 17(1), 145-168, Ocak 2024
[Online]: http://dergipark.org.tr/akukeg
DOI number: http://doi.org/10.30831/akukeg.1316111

Comparing the Feedback of University Supervisor and Cooperating Teachers for Preservice Science Teachers within the Scope of Pedagogical Content Knowledge

Fen Bilimleri Öğretmen Adaylarına Uygulama Öğretmenleri ve Uygulama Öğretim Elemanı tarafından verilen Geribildirimlerin Pedagojik Alan Bilgisi Kapsamında Karşılaştırılması

Işık Saliha KARAL EYÜBOĞLU*

Received: 17 June 2023 **Research Article** **Accepted:** 17 January 2024

ABSTRACT: To comprehend the perspectives of school and university mentors in evaluating practicum experiences based on educational reform, this study aimed to investigate the feedback, within the framework of pedagogical content knowledge, provided by two cooperating teachers and a university supervisor to preservice teachers. Observation and document were used for data collection. Data collected within the context of the Teaching Practice course were analyzed deductively, considering the components of pedagogical content knowledge. Findings showed that although feedback given by mentors was positive or corrective, specific changes suggested by mentors about the preservice teachers' teaching practices were rather corrective. Although mentors' feedback seems to be concentrated in the fields of 'instructional strategies and representation' and 'pedagogy,' the majority of feedback provided in other PCK components points out those mentors focused on student learning within the knowledge of 'student understanding'. Feedback indicated that mentors especially attached importance to students' understanding of the subject and active participation in the learning, which are the goals of a science teaching program. The other aims of teaching based on inquiry and students' self-responsibilities in learning seemed to be in the background. Although mentors' feedback pointed out some differences in the categories of curriculum and assessment, they were consistent in general.

Keywords: Feedback, pedagogical content knowledge, university supervisor, cooperating teacher, science, preservice teacher, teaching practice.

ÖZ: Reformlara dayalı öğretmenlik uygulamalarını değerlendirmede okul ve üniversite danışmanlarının bakış açılarını anlamak amacıyla bu çalışma ile iki uygulama öğretmeni ve bir uygulama öğretim elemanının fen bilimleri öğretmen adaylarına pedagojik alan bilgisi kapsamında verdikleri geribildirimlerkarşılaştırılmıştır. Gözlem ve doküman veri toplama araçları olarak kullanılmıştır. Öğretmenlik uygulaması dersi kapsamında toplanan veriler pedagojik alan bilgisi (PAB) bileşenleri dikkate alınarak tümden gelimli içerik analizi ile çözümlenmiştir. Bulgular danışmanların hem negatif (düzeltici) hem de pozitif geri bildirimler vermelerine karşın, adayların öğretimlerine ilişkin detaylı önerilerinin çoğunlukla düzeltici yönde olduğunu göstermiştir. Danışmanların geri bildirimleri 'öğretim yöntemleri ve sunum' ile 'pedagoji' bileşenlerine odaklanmış gibi görünse de, tüm kategorilerdeki geri bildirimlerin 'öğrenci anlaması' bileşeni çerçevesindeki öğrenci öğrenmesine odaklandığı sonucuna varılmıştır. Geri bildirimlerin fen öğretim programının amaçlarından özellikle öğrencilerin konuyu anlamaları ve öğretim sürecine aktif olarak katılmaları ile ilgili olduğu, sorgulamaya dayalı öğretim ve öğrencilerin kendi öğrenmelerinden sorumlu olma amaçlarının geri planda kaldığı görünmektedir. Danışmanların geri bildirimleri, program ve değerlendirme bileşenleri için bazı farklılıklar gösterse de genel olarak uyumlu görünmektedir.

Anahtar kelimeler: Geribildirim, pedagojik alan bilgisi, uygulama öğretim elemanı, uygulama öğretmeni, fen, öğretmen adayı, öğretmenlik uygulaması.

* Assoc. Prof. Dr., Giresun University, Giresun, Türkiye, saliha.karal@giresun.edu.tr, https://orcid.org/0000-0002-6966-9947

Citation Information
Karal Eyüboğlu, I. S. (2024). Comparing the feedback of university supervisor and cooperating teachers for preservice science teachers within the scope of pedagogical content knowledge. *Kuramsal Eğitimbilim Dergisi [Journal of Theoretical Educational Science], 17*(1), 145-168.

ISSN: 1308-1659

Developments in science and technology, changing needs of individuals and society and improvements in learning/teaching approaches require changes in expectations from students, preservice teachers, teachers, and teacher educators. Students are expected to grow up as individuals who can use information functionally, solve problems, and inquire. The science curriculum requires primary school teachers to ensure the active participation of students in the learning process as they are responsible for their own learning and to use learning strategies based on inquiry and knowledge transfer (Ministry of National Education [MoNE], 2018a). The reflection of reforms/innovations in the field of education to the learning environment is closely related to the qualifications and competencies of the teachers who guide this process. Teachers who will train students with these qualifications have to design the education process effectively and have professional skills as well as deep knowledge in their field. MoNE (2017), which has the feature of reference text in the regulation of the curriculum of higher education institutions that train preservice teachers in Turkey, in the preservice teacher training process and in the candidacy processes such as teaching practice, is defined as 'professional knowledge', 'professional skills' and 'attitudes and values'.

Pedagogical content knowledge (PCK) in the field of professional knowledge is seen as a teacher's ability to organize the concept or the subject according to the characteristics of learners and learning environments. PCK was described by Shulman (1987) as a special amalgam of content and pedagogy, and his definition includes two components: knowledge of instructional strategies and knowledge of students' understanding of the topics. This classification was extended by various research with other components, i.e., pedagogical knowledge, curriculum knowledge, and assessment knowledge (Jing-Jing, 2014). According to general competencies for teaching profession (MoNE, 2017), teachers should have professional knowledge and exhibit the knowledge and skills to deliver instruction effectively.

It is expected that preservice teachers (PTs) would develop and integrate their knowledge, skills, and attitudes gained in theoretical courses in the teacher training programs. The practicum field experience in teacher education is one of the major steps and influential factors in the preparation of preservice teachers (Badger, 2012; Steadman & Brown, 2011; Vertemara & Flushman, 2017). It gives PTs an opportunity to apply the theoretical concepts learned in the university classroom to the primary/secondary school classroom (Eck & Ramsey, 2019) and establish collaborative contexts for interactions among preservice teachers (PTs), cooperative teachers (CTs) and university supervisors (USs) who are putting into practice the ways of thinking, doing, and speaking advocated by reform documents (Van Zee et al., 2003). Throughout this experience, the PTs interact with the CT and the US, forming a cooperative triad that has a corrective role in implementing educational reform (Asplin & Marks, 2013). PTs are in a position to function as agents of reform, and mentors have a noticeable role in determining whether novices enact desired reform-based teaching practices and help spread these practices in their schools (Davis et al., 2006; Koballa & Bradbury, 2012). Because mentors are especially important in helping PTs to reflect upon their teaching and in providing access to a range of knowledge areas to assist PTs in their professional development, mentoring has come to be viewed as a means of reforming science teaching (Clarke et al., 2014; Koballa & Bradbury, 2012; Sandvik et al., 2019). USs

who serve as mentors share the responsibility of guiding PTs with CTs about reform-based science teaching and creating a connection between the practice schools and the university by visiting the schools (Vertemara & Flushman, 2017). They are expected to support PTs' implementation of theories learned in coursework with appropriate feedback to provide and enrich learning and skill development opportunities in the final stage of teacher preparation (Fernandez & Erbilgin, 2009; Hudson, 2014; Steadman & Brown, 2011). Although each group in the cooperative learning triad has a distinctive and complementary role to play in the teaching experience, the working relationship between them should be a close one that embodies parity of esteem and respects the contributions of all the partners (González-Toro et al., 2020). However, the practicum field experience is generally criticized (Grudnoff, 2011; Wilson, 2006) about the lack of partnership contacts between CTs and USs (Portelance et al., 2016; Sim, 2010), how theory and practice are reconciled in field experience (Allen et al., 2010; MacDougall et al., 2013) and a relative consensus between USs and CTs over what constitutes proper performance, professionalism and practice for maintaining professional consistency (Darling-Hammond, 2006). For example, even if PTs enter the classrooms with reform-based ideas about teaching and guided by CTs who rather value traditional notions of science teaching, the guidance constrains innovation and shapes the new teacher to fit the norms of the school (Bradbury & Koballa, 2007; Wilson, 2006). A few of the reasons for this are that PTs believe that classroom teachers have more realistic experience than their education faculty tutors, who are viewed as inspectors rather than collaborative partners (Asplin & Marks, 2013). For these reasons, it can be said that science teaching orientations of PTs affected shaped in the direction of reform-based approaches in teacher training programs are open to changes under the influence of CTs not meeting the expectations of reform-based science curriculum (Bates & Burbank, 2008; Hanuscin et al., 2011). On the other hand, it is emphasized that the role of the CT within the triad is secondary because the university is the final authority in the PTs' success (Clarke et al., 2014; Van Zee et al., 2003). Although the roles CTs and USs are distinct and/or changing, they are complementary in that the combination of their respective specific characteristics allows for consistency in student training (Burns et al., 2016; Portelance et al., 2016). For this, in a practicum experience, professional consistency is important for USs and CTs to share a relative consensus over what constitutes proper performance, professionalism, and practice (Darling-Hammond, 2006; Smith & Lev-Ari, 2005).

Regulation on the Teaching Practice of PTs in Educational Institutions affiliated with the Ministry of National Education requires each PT to receive teaching practice at the appointed school under the guidance of an experienced teacher by means of 'Teaching Practice' courses taken. When the Teaching Practice course is completed, PTs will be able to reach the competencies of the teaching profession by teaching classes with various levels in the practice school. PTs are evaluated at least four times each semester under the supervision of the CT and the instructor (MoNE, 2018b). The number of PT must be a maximum of four for school teachers and a maximum of eight for instructors. The CT evaluates the performance of the PTs in teaching experience with a weight of 70% and by the US with a weight of 30%. Although this regulation aimed to obtain coordination between CT, US, and PT in practice school activities, research studies on teaching experience in Turkey pointed out a lack of communication

and cooperation between the CTs and USs (Arkün-Kocadere & Askar, 2013; Polat et al., 2020; Topkaya et al., 2012) and PTs' concerns arising from gaps between supervisors and mentors' expectations (Paker, 2008). Inconsistency between the CTs and the USs and multiple perspectives suggested during the teaching experience by mentors can lead to conflicting messages for the PTs (Hudson, 2016; Tillema, 2009; Yayli, 2008). It is clear that to reinforce the reform-based teaching practices espoused in the preservice program for the school context, there should be close partnerships between universities and schools (Bradbury, 2010). Despite the emphasis on the need for harmonization between the interventions of CTs and USs (Portelance et al., 2016), certain obstacles, such as summative examinations consisting of memory recall, are inconsistent with learner-centred education with its origins in constructivism advocated by changed science curricula causes inconsistencies between them (Hume & Coll, 2007; Ranade, 2008). However, even if educational reforms match assessment methods, adequate teacher professional development is required to implement new curricula (Coll & Taylor, 2012). For instance, CTs need to know how to continue with the teaching practice in the new curriculum, i.e., what learner-centered education actually means in terms of teaching practice. Becoming a CT in the triad partnership provides an opportunity to reach new knowledge and tenets of reform-based science teaching as a result of interaction with faculty tutors (Clarke et al. 2014). This is an important aspect of the professional development of experienced science teachers who agree to serve as CTs but who might not be well-versed in the tenets of reform-based science teaching (Koballa & Bradbury, 2012).

One of the most effective ways to investigate the efforts of mentors to promote reform-based science teaching would be by examining feedback from CTs and USs following a teaching episode (Burbank et al., 2016; Sim, 2010; Tarekegn et al., 2020). PTs value feedback from CTs and USs within the frame of practicum field experience because they contribute to their perception of instruction, subject matter, and student learning and may affect a PT's decision to change or develop a practice (Hudson, 2016; Smith & Lev-Ari, 2005). With effective scaffolding and feedback, PTs can move from simplistic perspectives about the causes of classroom events to more expert understandings of how aspects of teaching and student development influence learning (Badger, 2012). Performance-based feedback to PTs is especially directly related to observed actions, and they are effective in practicum field experience, which improves teaching activities (Cornelius & Nagro, 2014). In spite of feedback providing is claimed to be useful in developing the teaching skills of PTs and making them competent (Chawla & Thukral, 2011; Hattie & Timperley, 2007; Tarekegn et al., 2020), some USs and CTs participating in mentoring did not supply enough feedback (Polat et al., 2020; Saka, 2019). Studies on feedback from USs and CTs (Nguyen, 2009; González-Toro et al., 2020; Tillema, 2009; Puttick & Wynn, 2020; Won et al., 2019) emphasized that the importance of training USs and CTs on how to effectively provide feedback. For example, Nguyen (2009) showed that the triad members were able to create a supportive environment when they communicated their areas of strengths and improvement to preservice teachers in a timely manner. Because immediate feedback reduces the practice of errors and provides correction before it is forgotten (Scheeler, 2008), examining feedback from school-based and university-based mentors concomitantly in teaching placement suggests that collect data of various triad meetings to more deeply

examine the giving/receiving of feedback by each member (Won et al., 2019). In this scope, while some studies focused on the feedback of cooperating teachers (Eck & Ramsey, 2019; Gurl, 2019; Jones et al., 2014; Matsko et al., 2020), others on university supervisors (Asplin & Marks, 2013; Bunton et al., 2002; Holbrook, 2022; Kastberg et al., 2020; Ritter et al., 2011). Research focusing on mentors' feedback pointed out that differences in mentor feedback can be a mismatch in mentors' expectations (Bradbury & Koballa, 2007; Hudson, 2014, 2016; Soares & Lock, 2007; Tillema, 2009). For instance, Soares and Lock (2007) demonstrated the differences in feedback provided by supervisors, with classroom management as a stronger focus than content knowledge or references to the lesson objectives. Kahan et al. (2003) compared the feedback profiles of CTs supervising the same PT and revealed different reasons for divergent feedback profiles. Because feedback influencing PTs' professional development reflects mentors' ability to review lesson plans, observe teaching, and provide constructive criticism about the teaching process (Hudson et al., 2005; Tarekegn et al., 2020), investigating the lesson observation feedback will give insight into the perspectives of CTs and USs. Therefore, this study aimed to investigate what feedback based on observation from CTs and the US provided for PTs to understand the perspectives of school and university-based mentors. Therefore, answers were sought for the following research questions:

1. What kind of feedback within the scope of PCK components did cooperating teachers and university supervisors provide to science preservice teachers?
2. To what extent is the feedback provided by cooperating teachers and university supervisors consistent?

Method

Because the aim of qualitative research is to examine natural environments without any special arrangement (Patton, 2014), in this work, a case study is adopted where an event is examined within its borders without any external interference and related behaviours (Yin, 2003).

Setting and Participants

The study was conducted in collaboration with two CTs in a state primary practice school and a US from the Faculty of Education. The CTs were assigned by the school administration, and the US was the mentor responsible for six PSTs who attended the Teaching Practice course of this practice school in the last semester of the training program. Because the US guiding PSTs were appointed by Faculty management and CTs were assigned by the school administration, the participants were selected using the convenience sampling method. One of the cooperating teachers, CT1, graduated from the faculty of education and had 12 years of teaching experience in state schools. The other, CT2, who completed his master's degree in science education and continued his Ph.D. studies in biology education, had 11 years of teaching experience in state schools. The US, who was also the researcher of this study, possesses 12 years of teaching experience in various state schools. This includes roles as a physics teacher in high and vocational high schools, a science teacher in primary schools, and five years of experience in the faculty of education. The PSTs in group CT1 were PST1, PST2, and

PST3, and those in group CT2 were PST4, PST5, and PST6. The teaching practices of the PSTs occurred at CTs' lecturing sessions and 5th, 6th, 7th, and 8th grade levels.

Data Collection

Observation and document were used as data collection instruments. The researcher, US, and two CTs observed the teaching practices of PSTs simultaneously during the lesson. Although the Faculty of Education recommended the use of structured observation forms for mentors, including some competencies such as presentation skills, subject matter, use of technology, and diversity in teaching, in this study unstructured observations was preferred to obtain flexibility for participants to emphasize the professional knowledge areas such as subject matter, pedagogical, curriculum. The participants took field notes during their observations and gave comments and feedback on PSTs' instructions. Each PST's instruction was observed for two lesson hours by the responsible CT and US. Table 1 shows the subjects taught and the teaching styles of PSTs.

Table 1

Observed Teaching Styles and Subjects of PSTs

PSTs	Subjects	Description of Teaching
PST1	Central nervous system	She started a discussion on a case using a question-answer method, presented the subject via lecturing, had students do an activity of creating a nervous system model in the elaboration, and used a worksheet in evaluation.
	Reproduction and growth in animals	She made an introduction by giving an example from life, had students play a game on amphigenesis in the exploration stage, presented the subject via lecturing, had students use drama in the elaboration, and requested students to write a related poem in evaluation.
PST2	Sense organs	In the beginning, she had a few students perform an activity, started a discussion on this activity in the exploration, explained the subject via lecturing, played a video in the elaboration, and organised an instructional game in evaluation.
	Refraction of light	She made an introduction using a material, had students experiment with the exploration, discussed the results with students, presented a video in the elaboration, and requested students to write a related poem in evaluation.
PST3	Lenses	She had an introduction with examples from life, did an experiment first by herself and then with groups of students, discussed the data and results with students, explained the subject, presented a video in the elaboration, and used an instructional game in evaluation.
PST4	Refraction of light	He made an introduction with question-answers, had a group of students perform experiments, requested students to develop arguments about results and make discussion, explained the subject via video, made a demonstration experiment in the elaboration, and used worksheets in evaluation.
	Electric Circuit Elements	Following the solution of the questions in the worksheet together with students, he continued answering the evaluation questions in the computer environment with students.

PST5	Bio-diversity	He started the lesson with question-answers and discussion, activated an instructional game in the exploration, and presented the subject via lecturing. The lesson expired.
	Growth and Development of Plants	He started the lesson with question-answers. He did a demonstration experiment on germination in the exploration but was not able to complete it due to failure in planning and shifted to the smart board for the explanation, elaboration, and evaluation cycles.
PST6	Agamogenesis and Amphigenesis	He started by exposing a model of a flower, allowing student groups to explore the model, explaining the subject with examples, having students use drama for pollination in the elaboration, and urging students to solve a puzzle in evaluation.

The other data collection instrument was the documents of field notes, including feedback provided to PSTs by the CTs and the US. Field notes consisting of two parts, descriptive and reflective information taken during observations, are widely recommended in qualitative research as a means of documenting the needed contextual information (Phillippi & Lauderdale, 2018). Field notes mentioned in this study were records of PTs' activities and their evaluation by the CTs and the US. Because CTs here generally preferred giving verbal feedback to PTs, the researcher requested the CTs give written feedback following lesson observation (Kastberg et al., 2020; Puttick & Wynn, 2020; Schwartz et al., 2018).

Data Analysis

Whether a deductive or an inductive approach is used in the analyses of data depends on the research questions and the general aim of the study (Elo & Kyngäs, 2008). Inductive approaches are often used when there is little knowledge about the phenomenon, while deductive approaches are used on the basis of previous studies and knowledge. In this study, data were analysed with deductive coding. Firstly, the statements in documents 'she talked to a certain group of students, other students were left on her backside causing some feeling of distraction' or 'while solving problems on the blackboard let us talk to the whole class, not to a few students' was coded as 'concentrating on specific students' and therefore were labelled as positive or corrective feedback. Positive feedback (PF) increases supervisees' confidence by pointing out their knowledge and skills, thus contributing to the competence perceived by those supervisees (Komiskey & Hulse-Killacky, 2004). Corrective feedback (CF), sometimes referred to as negative, is a term that clearly indicates a desire for a specific change in the student teachers' practice (Bjørndal, 2020). Then, these codes were associated with professional knowledge categories, i.e., subject matter knowledge, pedagogical knowledge, and curriculum knowledge. Through these codes, six professional knowledge categories corresponding to components of PCK were generated. For example, 'selection of efficient and appropriate activities (IS14)' was categorized in instructional strategies and representation, and 'classroom management (P5)' was in pedagogy, as seen in Table 2. Table 2 shows the six PCK components: subject matter, pedagogy, assessment, curriculum, student understanding, and instructional strategies.

Table 2

PCK Components and Feedback Codes

Subject Matter	SM_1	Giving incorrect knowledge	Instructional Strategies and Representation	IS_1	Preparing an effective lesson plan
	SM_2	Being unable to answer student's question or fudge		IS_2	Teaching in accordance with the lesson plan
	SM_3	Confusion of concepts		IS_3	Using educational technology
	SM_4	Satisfactory subject matter knowledge		IS_4	Using the course book
	SM_5	Deficient concept mapping		IS_5	Using the blackboard effectively
Pedagogy	P_1	Tone of voice		IS_6	Summarizing the subject
	P_2	Calling the student by name		IS_7	Daily examples
	P_3	Concentration on specific students		IS_8	Organizing various activities
	P_4	Standing at a specific place in the classroom		IS_9	Effective use of visuals or videos
	P_5	Classroom management		IS_{10}	Doing or promoting experiment
	P_6	Monotone speech		IS_{11}	Presenting the subject in a prescribed time period
	P_7	Consistency in behaviour		IS_{12}	Having students take notes
	P_8	Walking around the class		IS_{13}	Clear activity/experiment directives
	P_9	Being fair to students		IS_{14}	Selection of efficient and appropriate activities
	P_{10}	Turning back to the class		IS_{15}	Group working
Student Understanding	SU_1	Present the subject fast	Assessment	A_1	Preparation of materials for assessment
	SU_2	Speaking fast		A_2	Using materials for assessment
	SU_3	Emphasizing important points		A_3	Asking for information, not commenting during teaching
	SU_4	Examining students' prior knowledge		A_4	Involving all students in the assessment
	SU_5	Giving enough time to students		A_5	Suitable assessment
	SU_6	Giving contradictory knowledge			
	SU_7	Present concepts not included in the subject	Curriculum	C_1	Lesson plan incompatible with curriculum
	SU_8	Presenting the subject in the correct order		C_2	Teaching incompatible with curriculum
	SU_9	Giving explanation during the examination of the prior knowledge			

Role of the Researcher

The researcher is the US who is one of the participants in this study, and she was the complete participant contributed to the internal validity by taking on the role of an insider, becoming a member of the group being studied, and spending a sufficient but not too long to cause bias a time with PSTs (Christensen & Johnson, 2004). Because she is a member of the natural environment, it is believed that the effects of the researcher's existence are limited. At the same time, her long-term experience as a science teacher in public schools contributed to her communicating effectively with CTs.

Researchers have taken some measures to ensure the trustworthiness of this study. Observations were made in different parts of the classroom so that the CTs and the US could independently reflect their own interpretations without being influenced by each other's thoughts. To mitigate the interpretive bias of a single researcher, the analysis of data was started after all data were gathered (McAlister et al., 2017). The researcher returned to the data at other times for intracoder reliability, which refers to consistency in how the same person codes data at multiple time points and transparency in the process of coding and creating thematic structures (O'Connor & Joffe, 2020).

Ethical Procedures

The search was approved by the Ethics Committee of Giresun University (Approval No: 2021/14-23).

Results

Findings on mentors' feedback are presented below as two subsections titled 'Feedback from the Cooperating Teacher 1 and the University Supervisor' and 'Feedback from the Cooperating Teacher 2 and the University Supervisor'. Table 3 and Table 4 show the feedback given by CTs and the US to all PSTs and their categories of PCK components.

Feedback from the Cooperating Teacher 1 and the University Supervisor

In this section, feedback from the cooperating teacher 1 (CT1) and the university supervisor (US) for preservice teachers (PST1, PST2, and PST3) are presented.

Table 3

Feedback of CT1 and US to PST1, PST2, PST3 and Related PCK Components

PST	Category of PCK Component	Feedback Type of Cooperating Teacher 1		Feedback Type of University Supervisor		Feedback Frequency
		Positive	Corrective	Positive	Corrective	
PST1	Subject Matter		SM_1, SM_5		SM_2	3
	Pedagogy	P_1, P_2	P_3, P_4, P_8	P_1, P_2, P_3	P_4, P_5	10
	Instructional Strategies and Representation	IS_1, IS_2, IS_4, IS_6, IS_8	IS_3	IS_1, IS_2, IS_4, IS_6, IS_5, IS_7	IS_3	13
	Student Understanding		SU_1, SU_3, SU_4		SU_1, SU_2	5
	Assessment	A_1	A_5	A_2		3
PST2	Subject Matter		SM_2, SM_3		SM_1, SM_2	4
	Pedagogy	P_1, P_2	P_3, P_5, P_6	P_1, P_2	P_3, P_4, P_5	10
	Instructional Strategies and Representation	IS_1, IS_8	IS_{13}	IS_2, IS_5, IS_6, IS_8, IS_9	IS_3, IS_{13}	10
	Student Understanding		SU_4			1
	Assessment		A_3	A_1		2
PST3	Subject Matter	SM_4		SM_4	SM_2	3
	Pedagogy	P_1, P_2, P_3, P_5, P_9	P_4	P_1	P_2, P_4, P_5, P_7	11
	Instructional Strategies and Representation	IS_1, IS_2, IS_9, IS_{10}	IS_{11}, IS_{12}	IS_1, IS_2, IS_3, IS_6, IS_8, IS_{10}	IS_5, IS_{13}	14
	Student Understanding		SU_5		SU_5	2
	Assessment	A_1, A_4		A_1		3
Total Feedback Frequency		24	22	27	21	

According to Table 3, the feedback of CT1 and US was positive and corrective in type. Although the number of feedback given by the two mentors was the same on average, the number of positive feedback supplied by US was higher than that of CT1. It is seen that feedback from the supervisors was concentrated in the categories ‘instructional strategies and representation’ and ‘pedagogy.’

The Category of Instructional Strategies and Representation

In this category, while CT1 gave positive feedback to PSTs on preparing an effective lesson plan, teaching according to the lesson plan, using the course book, summarizing the subject, effective use of visuals, and doing the experiment, the US

gave positive feedback (PF) also on using educational technology and blackboard effectively, talking on daily examples and organizing various activities:

> At the end of an inquiry-based experiment carried out on the refraction of light, PST2 wrote the comments on observations of each group on the blackboard and compared the results. She explained and summarized the correct result attained (US, PF, IS_6).
>
> PST1 used the blackboard and the textbook; it was positive (CT1, PF IS_4, IS_5).

CT1 and US gave the same corrective feedback on using educational technology and clear activity/experiment directives:

> For PST1, using computer-stored figures would be better since she was not good at drawing figures. Unrealistic figures may lead to incorrect learning for students (US, CF, IS_3).
>
> PST2, by giving deficient information about how the presentation and activities would be done, hindered students' effective participation in and enjoyment of the activity (CT1, CF, IS_{13}).

Different corrective feedback provided by the supervisors were on presenting the subject in the prescribed time period, having students take notes during lessons, and using the blackboard effectively:

> PST3 did not have students take notes in their notebooks (CT, CF, IS_{12}).
>
> PST3 drew three different figures of lenses side by side and asked students what kind of lens each figure represents. She did not name the figures as 1, 2, or 3, so the students had to refer to them as 'this,' 'that,' etc., causing confusion (US, CF, IS_5).

The richest feedback supplied by mentors was seen in the category 'instructional strategies and representation,' and the category of pedagogy followed this.

The Category of Pedagogy

The feedback in the category of pedagogy was concerned with monotone speech, tone of voice, calling the student by name, being interested in specific students, standing at a specific region in the classroom, classroom management, consistency in the behavior, walking around the classroom and being fair to students. Positive feedback of CT1 and US was mostly related to the tone of voice and calling students by name:

> Tone of voice of PST2 was good (CT1, PF, P_2)

The positive feedback given by CT1 in this category, who gave more positive feedback than the instructor, was related to 'concentration on specific students' and 'being fair to students':

> PST3 gave students the right to speak as equally as possible (CT1, PF, P_9)

Supervisors gave corrective feedback to PSTs, especially about concentration on specific students, standing at a specific place in the classroom, and classroom management:

> While PST2 talked to a certain group of students, other students were left on her backside, causing some feeling of distraction (CT1, CF, P_3).
>
> PST1 spent much of her time near the table and the blackboard without walking in the classroom (US, CF, P_4).

Feedback from CT1 and the US to PST2 and PST3 about clear activity/experiment directories in the instructional strategies and representation category are also related to pedagogy. US emphasized that the reason for the trouble PST3, who was generally successful in classroom management in lab activities, was that she did not provide clear experimental directives:

> Before starting an activity… you can explain each stage of it and write down some steps… Thus, you don't have to repeat the same explanation to every group; also, other students will not be idle when you're busy with a group (US, CF, IS_{13}).

The Category of Subject Matter

The feedback in the category of subject matter was about giving incorrect knowledge, being unable to answer student's question or fudge, confusion of concepts and satisfactory subject matter knowledge. All feedback in the category 'subject matter' to PSTs, except PST3, was corrective because mentors reported that PSTs did not have rich subject matter knowledge:

> PST2 was not able to satisfactorily answer the student's question, 'where is the eardrum exactly?'. Similarly, she said, 'we would have seen objects in two dimensional if we had only one eye' (US, CF, SM_1, and SM_2).
>
> PST1 made some explanations that would cause misunderstandings among students. For example, she said that the bat was a bird. A bat is a mammal. She had difficulty in answering students' unexpected questions because she probably learned by reciting (CT1, CF, SM_1).

CT1 and US gave the same positive feedback for PST3:

> Subject matter knowledge of PST3 is satisfactory (US, PF, SM_4).

The Category of Assessment

Mentors gave a small amount of feedback in the 'assessment' and 'student understanding' categories. While the US focused on preparation and use of materials for assessment and only gave positive feedback in the category 'assessment,' CT1 added feedback about 'involving all students in assessment.' The corrective feedback in the category of 'assessment' was supplied only by CT1:

> PST1 prepared a worksheet for evaluation, but the questions in the worksheet are of lower level. She had to ask selective and specific questions (CT1, PF, CF, A_5).

The Category of Student Understanding

All feedback in the student understanding category was corrective, and two mentors supplied similar corrective feedback to the PTSs. For example, both of them stated that quick presentation of PST1 affected student understanding negatively:

> Presentation of the subject has to possess integrity and hierarchy. PST1 shifted from one concept to another, and this created confusion in student's minds (CF1, CF, SU_1).
>
> In the presentation of the subject, a procedure from simple to complex should be followed. PST1 implemented his lesson plan and prepared for two hours within one hour; she was speedy (US, CF, SU_1).

In sum, the amounts of positive and negative feedback of CT1 and US are close to each other; the largest amount of feedback was in the category of instructional strategies, and the majority of this feedback was positive. Mentors especially appreciated PTs for using visuals and conducting activities and experiments in which students were active. The second largest amount of feedback occurred in the category of pedagogy, and mentors supplied similar positive and corrective feedback for all PSTs, except for PST3. Almost all feedback, with small amounts in the categories of subject matter knowledge and student understanding, was corrective. While the US supplied completely positive feedback in the category of assessment, the CT gave corrective feedback as well.

Feedback from the Cooperating Teacher 2 and the University Supervisor

In this section, feedback from the cooperating teacher 2 (CT2) and the university supervisor (US) for preservice teachers (PST4, PST5, and PST6) are presented.

Table 4

Feedback of CT2 and US to PST4, PST5, PST6 and Related PCK Components

PST	Category of PCK Component	Feedback Type of Cooperating Teacher 2		Feedback Type of University Supervisor		Feedback Frequency
		Positive	Corrective	Positive	Corrective	
PST4	Subject Matter		SM_1, SM_2		SM_1, SM_2	4
	Pedagogy	P_2	P_3, P_7, P_9	P_1, P_2	P_2, P_4	8
	Instructional Strategies and Representation	IS_{15}	IS_1, IS_5, IS_6, IS_7, IS_{12}, IS_{14}	IS_8, IS_{10}	IS_1, IS_2, IS_3 IS_6, IS_9, IS_{13}	15
	Student Understanding		SU_7		SU_6	2
	Assessment	A_1		A_2	A_5	3
	Curriculum		C_2			1
PST5	Subject Matter		SM_4		SM_3	2
	Pedagogy		P_3, P_9, P_{10}	P_1, P_5	P_2, P_4	7
	Instructional Strategies and Representation	IS_{15}	IS_1, IS_2, IS_4, IS_5, IS_7, IS_{13}	IS_3, IS_5, IS_{11}, IS_{12}	IS_1, IS_2, IS_6, IS_9	15
	Student Understanding		SU_7		SU_2	2
	Assessment		A_1	A_1	A_5	3
	Curriculum		C_1, C_2			2
PST6	Subject Matter		SM_3		SM_3	2
	Pedagogy		P_3, P_5	P_{10}	P_1, P_2, P_5	6
	Instructional Strategies and Representation	IS_5, IS_{15}	IS_1, IS_9	IS_9, IS_5, IS_{11}, IS_{15}	IS_1, IS_2, IS_9, IS_{14}	12
	Student Understanding		SU_1, SU_3, SU_8	SU_4, SU_5	SU_1, SU_9	7
	Assessment		A_1, A_2	A_1		3
Total Feedback Frequency		6	37	20	31	

Table 4 shows that although feedback from CT2 and US was positive and corrective in type, the corrective feedback numbers of both mentors were greater than those of positive feedback. The amount of feedback supplied by the US was more than that of CT2, and it seems that this difference was due to the amount of positive feedback, with CT2 giving only six positives in total. It is seen that feedback from the

supervisors was concentrated in the category 'instructional strategies and representations.

The Category of Instructional Strategies and Representation

Feedback from the supervisors was concentrated in the category of 'instructional strategies and representation,' similar to Table 3. In this category, while CT2 supplied positive feedback on using the blackboard effectively and group working, US gave feedback on using visuals and technology, organizing various activities, doing or promoting experiments, group work, using the blackboard effectively, and having students take notes. The corrective feedback of US was on preparing an effective lesson plan, teaching according to the lesson plan, using educational technology, summarizing the subject, effective use of visuals, clear activity/experiment directives, and selecting efficient and appropriate activities:

> The drama used by PST6 to explain the parts of a flower was not a suitable instructional strategy (US, CF, IS_{14}).
>
> PST4 was not able to operate the smart board during teaching… he probably did not practice its usage before (US, CF, IS_{3}).

CT2 gave corrective feedback similar to those of US, except feedback on teaching according to the lesson plan and using educational technology:

> It would be better for PST5 to write on the blackboard what each group should do during the activity; disruption happened, and the students were not able to understand (CT2, CF, IS_{13}).
>
> PST4 did not choose an activity complying with the subject (CT2, CF, IS_{14}).

But CT2 also gave additional corrective feedback on using the course book and blackboard, daily topics, and having students take notes:

> It would be better for PST2 to write on the board what the group would do in the activity (CT2, CF, IS_{13}).

Supervisors appreciated PSTs for planning and carrying out various activities:

> Group working activity of PST4 was good (CT2, PF, IS_{15}).
>
> PST6 used a flower model and a video in his explanations (US, PF, IS_{9}).

However, they gave corrective feedback emphasizing that the final results following discussion and observations were not clarified well:

> PST4 used empty and water-filled glasses and water mixed with vinegar in the experiment and asked students to observe the appearance of the fork. The students explained their ideas, but the result was not clearly stated… He did not present a clear summary (US, CF, IS_{6}).
>
> PST4 did not give any explanation following the activity. What was the reason why light is refracted differently in water, vinegar, and air? An explanation should be made because concluding a result following the experiment is not easy for every student (CT2, CF, IS_{6}).

The Category of Pedagogy

The category of pedagogy follows, in frequency, that of instructional strategies. Although CT2 gave positive feedback on PST4's about calling the student by the name, US gave positive feedback for every PT on tone of voice, calling the student by the name and turning back to the class:

> PST4 knew students and called them by their names (CT2, PF, P_{2}).

While the corrective feedback of US concentrated on calling the students by name and standing at a specific place in the classroom, the feedback of CT2 paid

attention to the code's concentration on specific students, consistency in the behavior, fair treatment to students and turning back to the class:

> PST6' tone of voice is too low; the speech is not heard and understood, which weakens his classroom management (US, CF, P_1).
>
> PST5, while solving problems on the blackboard, let us talk to the whole class, not to a few students. Otherwise, students will become busy with other businesses (CT2, CF, P_3).
>
> PST4 has to be consistent and fair in the class. Fair treatment is vital; otherwise, the teacher will lose esteem (CT2, CF, P_9).

The Category of Student Understanding

The third, in frequency order, category includes feedback essentially on student understanding. Most of this feedback was corrective and on presentation or speaking in a fast mood:

> PST6 presented the subject rather quickly… (CT2, CF, SU_8).
>
> PST6 presented all concepts in a furry in the first session and was short of teaching material for the second session, so he had to make repetitions (US, CF, SU_1).

While corrective feedback supplied by the CT2 was on emphasizing important points, presenting concepts not included in the subject, and presenting the subject in the correct order, US concentrated on giving contradictory knowledge and giving explanations during the examination of the prior knowledge:

> PST6 submitted some information before examining the prior knowledge of students on the subject (US, CF, SU_9).

The Category of Subject Matter

Feedback in the category of subject matter about all PTs was corrective, perhaps because supervisors wrote that PTs did not have rich subject matter knowledge:

> PST4 answered the student's question 'what is the distance from the earth to the sun' as '8 light years', indicating that his subject matter knowledge was poor (CT2, CF, SM_1).
>
> PST4 said to the student who claimed that the brightness increases with the number of cells the opposite, but later, he repeated what the student claimed to the class. His explanations were not consistent (US, SM_1).

The Category of Assessment

In the assessment category, supervisors paid attention to the preparation and use of materials for assessment. CT2 stated that PSTs except PST4 did not carry out the evaluation, and he appreciated the assessment activities of PST4:

> The evaluation activity (worksheet) chosen by PST4 was quite good (CT2, PF, A_1).
>
> PSTs generally did not assess students' understanding; if they had some time left, they asked questions written at the end of the chapter; they did not take students' questions into account because they focused mainly on presenting the subject (CT2, CF, A_1).

Similarly, US gave positive feedback to the PSTs about preparing and using assessment activities and also focused on the quality of the assessment questions used by them:

> PST5 prepared an assessment worksheet and used it in the classroom, but the selected questions were not suitable (US, CF, A_5).

The Category of Curriculum

Feedback in the category of curriculum was supplied only by CT2, drawing attention to inconsistency between the lesson plans/activities and curriculum targets for 2 of 3 PSTs:

> PST4's activity, in which laser light illuminates the water stream flowing through the hole at the side surface of a bottle, is related to total internal reflection. The students have not learned this subject yet (CT2, CF, C_2).
>
> PST5 talked about concepts, such as dormancy and anaerobic respiration, not included in the learning objectives, thus causing confusion in student's minds (CT2, CF, SU_7, C_2).

In sum, there were noticeable differences between the numbers of positive and corrective feedback of CT2 and US. The number of corrective feedback from both mentors was more than their positive feedback, and only 15% of the CT2's feedback was positive. The largest number of feedback was in the category of instructional strategies, and the majority of this feedback was corrective. While CT2 appreciated PTs only for the group working in this category, the US gave a number of positive feedback. All feedback from mentors in the category of subject matter knowledge was corrective. Feedback from US for this PST group in the assessment category was positive and negative, similar to those of CT2. Although the feedback of CT2 in the pedagogy and student understanding categories was mostly corrective, those of the US were positive and corrective. Feedback in the curriculum category was supplied only by CT2 and was corrective.

Discussion and Conclusion

The first research question in this study was what kind of feedback CTs and US provided to PSTs. Findings showed that although feedback given by mentors was positive as well as corrective, the corrective ones were more than the positives, contrary to some others' studies (Bullough, 2005; Kahan et al., 2003), emphasizing that all CTs' feedback was more positive than corrective. This corrective feedback showed that mentors required specific changes in the PTs' teaching practices (Bjørndal, 2020). Because mentors providing feedback were also responsible for the final assessment of the teaching practice activities, giving corrective feedback may be considered to be challenging for both PTs and mentors (Bjørndal, 2020; Tang & Chow, 2007). However, in the teacher education context of this study, CTs and USs tended to give high final grades to PSTs for their teaching practices. Most USs do not participate in PTs' teaching because of claimed time restrictions or, as CTs asserted, of neglect (Andrew, 2007; Hellison, 2003; Topkaya et al., 2012). As a result, these types of USs either give high grades to all PSTs or leave the decision to the initiative of the CT, who becomes decisive in assessing (Weiss & Weiss, 2001). Another type of university supervisors who observe the teaching of PTs abstains from giving realistic assessment grades, which would create negative emotional reactions among PTs (De la Cruz et al., 2015). Similarly, CTs think that the number of PTs per CT is too high because they already have much work to do, thus providing limited teaching experience opportunities for PTs (Saka, 2019), so their assessments, although free-handed, may not be fair enough (Arkün-Kocadere & Askar, 2013). In this study, informal conversations about PSTs between CTs and US showed that if PSTs were rigorous and willing in the teaching experience process, CTs would not take deficiencies of PSTs into account in grading.

These circumstances may clarify why mentors did not worry about giving corrective feedback, a crucial part of mentoring in teaching placements (Amobi, 2005; Crasborn et al., 2008; Crutcher & Naseem, 2016). Although satisfactory final grades of the triad are generally expected, the corrective feedback items and dozes point out that mentors intended to improve PSTs teaching and enhance their thinking ability beyond teaching to analyse, reflect, and reconstruct their teaching (Range et al., 2013).

Mentors are ideally expected to provide feedback in both content-specific and general pedagogies (Schwartz et al., 2018) because PCK is an amalgam of content and pedagogical knowledge (Shulman, 1986) and closely related to 'the ways of representing the subject that make it comprehensible to others' (Shulman, 1987). The importance of representing the subject may explain the concentration in the categories of 'instructional strategies and representation' and 'pedagogy.' Similarly, Won et al. (2019) found that both CTs and USs mainly provided feedback on key areas such as student engagement, more effective use of instructional norms, and application of content-based pedagogies. Subject planning and presentation by PSTs provided insight to mentors not only about PSTs' knowledge of instructional strategies but also about other professional knowledge categories, i.e., subject matter, student understanding, curriculum, and assessment. Although feedback from the mentors seems to be concentrated in the fields of instructional strategies and pedagogy, the majority of feedback in all knowledge categories points out that the mentors focus on student understanding, meaning that the transfer of knowledge is important. In reform-based science curricula for elementary and secondary schools (MoNE, 2018), a holistic perspective has been adopted in terms of learning-teaching theories and practices based on knowledge transfer, inquiry, and active participation in the learning process where students are responsible for their own learning. Teachings of PSTs based on the constructivist approach showed that they made efforts to put into practice what they learned in theoretical courses about the inquiry approach. Feedback indicated that mentors especially attached importance to student's understanding of the subject and active participation. Issues on teaching based on inquiry and student's self-responsibilities in learning seemed to be in the background as Bradbury and Koballa (2007) reported that dialogues between CTs and PTs focused on general pedagogical knowledge instead of the nature of science, scientific inquiry, and literacy issues which are the central elements of reform in science teaching. Although mentors provided corrective feedback to PSTs helping novices match classroom practice with reform-based views of teaching, including an emphasis on inquiry (Bradbury, 2010), they neither criticized any PST for not particularly using inquiry teaching (Furtak et al., 2012) nor did they appreciated any other for practicing this sort of teaching. The mentors' statements about PSTs' inquiry teaching, such as 'concluding a result following the experiment is not easy for every student,' may give an idea about persistence in making and dictating a summary of observations, experiments, and discussions. CTs seemed to think that students did not have skills for interpreting information and drawing conclusions; thus, they may not prefer to provide feedback on the development of procedural and epistemic inquiry (Furtak et al., 2012).

The second research question asks to what extent feedback from the CKs and US are consistent. This research shows that mentors provided very specific feedback more effective in contributing to the improvement of PSTs' teaching than general classroom

practices (Getachew et al., 2020; Moore, 2003; Scheeler et al., 2004). Although mentors' feedback points out some differences in the categories of curriculum and assessment, they were consistent in general. Because CTs were more engaged with the curriculums and national exams than the USs (Chaliès et al., 2004), it can be considered that CTs gave more detailed feedback in these categories. Similarly, positive and corrective feedback from CTs concentrated on nearly the same professional knowledge area. One reason for the difference in the feedback of the two CTs was that PSTs under the guidance of CT1 and CT2 were different. The fact that the US gave more corrective feedback to PSTs under the guidance of the CT2 also seems to support this situation. It does not seem possible in this study to claim anything about other reasons for differences in the feedback of CTs (Hudson, 2014).

This study was conducted with a limited number of participants, revealing CT and US's feedback, which was consistent and similar, but the reason for this consistency was that the professional experience of US as a science teacher in primary schools was similar to CTs and gives in general science lectures with lab works not teaching methods courses in the teacher training program. It is clear that another study to be conducted with USs having different science teaching orientations may reveal different findings. As a matter of fact, feedback from US who conduct instructional teaching courses having an impact on PST's argument-based teaching plans would be mainly based on understanding the nature of science by designing investigations, collecting data, and using evidence to support findings through collaboration and discourse (Bradbury, 2010).

Another limitation of the study was that the feedback from the mentors was given during the initial teaching practices of the PSTs. Because feedback may affect the subsequent teaching of PSTs, the distribution of feedback given by mentors in categories may change, and feedback concentrated in the categories of instructional strategies and pedagogy may shift to different knowledge categories. Next, a longitudinal development research design will contribute to understanding how mentors' feedback changes throughout the Teaching Practice course.

The researcher's experiences indicated that CTs are open to cooperation and communication with university supervisors. They care about the thought of the US and want to give importance to the thought of themselves (Shantz, 1995). In this study, the most important reason for the effective/close relationship offering opportunities for sharing knowledge and skills between US and CTs (Allen et al., 2010) was the US' attitude due to the common corporate culture. Through empathic communication, thoughts CTs about problems as there is not enough information about the implementation of the inquiry-based learning approach and not enough examples and explanations for the learning outcomes (Bekmezci & Ateş, 2017) or CTs having little experience in planning and conducting inquiry-based science activities with their students and may be unwilling or unable to model that strategy for a novice can change.

Conflicts of Interest

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Author Bio:

Işık Saliha Karal Eyüboğlu works as an Associate Professor at the Giresun University, Faculty of Education, Department of Mathematics and Science Education. Working interests; teacher education, physics education.

References

Allen, J. M., Butler-Mader, C., & Smith, R. A. (2010). A fundamental partnership: the experiences of practising teachers as lecturers in a preservice teacher education programme. *Teachers and Teaching: Theory and Practice, 16*(5), 615-632.

Amobi, F. A. (2005). Preservice teachers' reflectivity on the sequence and consequences of teaching actions in a microteaching experience. *Teacher Education Quarterly, 32*(1), 115-126.

Andrew, L. (2007). A new university supervisor: Their experiences and enrichment. *Essays in Education, 22*(1), 12-27.

Arkün-Kocadere, S., & Askar, P. (2013). A review of views about student teaching courses and an application model proposal. *H.U. Journal of Education, 28*(2), 27-43.

Asplin, K. N., & Marks, M. J. (2013). Increasing the influence of university supervisors during student teaching. *The Professional Educator, 37*(1), 237-342.

Badger, J. (2012). Analyzing levels of feedback delivered by cooperating teachers and supervisors in a teacher internship: A case study. *Georgia Educational Researcher, 9*(1), 21-39.

Bates, A. J., & Burbank, M. D. (2008). Effective student teacher supervision in the era of no child left behind. *The Professional Educator, 32*(2), 1-11.

Bekmezci, S. M., & Ateş, Ö. (2017). Science curriculum from the perspectives of Turkish teachers: Problems encountered and suggestions for solutions. *European Journal of Education Studies, 3*(10), 137-158.

Bjørndal, C. R. (2020). Student teachers' responses to critical mentor feedback: A study of face-saving strategies in teaching placements. *Teaching and Teacher Education, 91*, 103047.

Bradbury, L. U. (2010). Educative mentoring: Promoting reform-based science teaching through mentoring relationships. *Science Education, 94*(6), 1049-1071.

Bradbury, L. U., & Koballa, T. R. (2007). Mentor advice giving in an alternative certification program for secondary science teaching: Opportunities and roadblocks in developing a knowledge base for teaching. *Journal of Science Teacher Education, 18*(2), 817-840.

Bullough, R. V. (2005). Being and becoming a mentor: School-based teacher educators and teacher educator identity. *Teaching and Teacher Education, 21*(2), 143-155.

Bunton, D., Stimpson, P., & Lopez-Real, F. (2002). University tutors' practicum observation notes: Format and content. *Mentoring and Tutoring: Partnership in Learning, 10*(3), 233-252.

Burbank, M. D., Bates, A., & Gupta, U. (2016). The influence of teacher development on secondary content area supervision among preservice teachers. *The Teacher Educator, 51*(1), 55-69.

Burns, R. W., Jacobs, J., & Yendol-Hoppey, D. (2016). The changing nature of the role of the university supervisor and function of preservice teacher supervision in an era of clinically-rich practice. *Action in Teacher Education, 38*(4), 410-425.

Chaliès, S., Ria, L., Bertone, S., Trohel, J., & Durand, M. (2004). Interactions between preservice and cooperating teachers and knowledge construction during post-lesson interviews. *Teaching and Teacher Education, 20*(8), 765-781.

Chawla, V., & Thukral, P. (2011). Effects of student feedback on teaching competence of student teachers: A microteaching experiment. *Contemporary Educational Technology, 2*(1), 77-87.

Christensen, B. J., & Johnson, B. (2004). *Educational research: Quantitative, qualitative, and mixed approaches, Research Edition.* Allyn & Bacon.

Clarke, A., Triggs, V., & Nielsen, W. (2014). Cooperating teacher participation in teacher education: A review of the literature. *Review of Educational Research, 84*(2), 163-202.

Coll, R. K., & Taylor, N. (2012). An international perspective on science curriculum development and implementation. In B. J. Fraser, K. G. Tobin, & C. McRobbie (Eds.), *Second international handbook of science education* (pp. 771-782). Springer.

Cornelius, K. E., & Nagro, S. A. (2014). Evaluating the evidence base of performance feedback in preservice special education teacher training. *Teacher Education and Special Education, 37*(2), 133-146.

Crasborn, F., Hennissen P., Brouwer, N., Korthagen, F., & Bergen, T. (2008). Promoting versatility in mentor teachers' use of supervisory skills. *Teaching and Teacher Education 24*(3), 499-514.

Crutcher, P. A., & Naseem, S. (2016). Cheerleading and cynicism of effective mentoring in current empirical research. *Educational Review, 68*(1), 40-55.

Darling-Hammond, L. (2006). Constructing 21st-century teacher education. *Journal of Teacher Education 57*(3), 300-314.

Davis, E. A., Petish, D., & Smithey, J. (2006). Challenges new science teacher's face. *Review of Educational Research, 76*(4), 607-651.

De la Cruz, M. S. D., Kopec, M. T., & Wimsatt, L. A. (2015). Resident perceptions of giving and receiving peer-to-peer feedback. *Journal of Graduate Medical Education, 7*(2), 208-213.

Eck, C. J., & Ramsey, J. W. (2019). An analysis of cooperating teacher feedback: a qualitative inquiry. *Journal of Research in Technical Careers, 3*(2), 97-113.

Elo, S., & Kyngäs, H. (2008). The qualitative content analysis process. *Journal of Advanced Nursing, 62*(1), 107-115.

Fernandez, M. L., & Erbilgin, E. (2009). Examining the supervision of mathematics student teachers through analysis of conference communications. *Educational Studies in Mathematics,* 72(1), 93-110.

Furtak, E. M., Seidel, T., Iverson, H., & Briggs, D. C. (2012). Experimental and quasi-experimental studies of inquiry-based science teaching: A meta-analysis. *Review of Educational Research, 82*(3), 300-329.

Getachew, T., Terfa, D., Tadesse, M., Atnafu, M., & Alemu, M. (2020). Ethiopian preservice primary science teachers' perceptions of mentoring in science teaching. *Journal of Science Teacher Education, 31*(8), 894-913.

González-Toro, C. M., Cherubini, J. M., Doig, S. R., & Fernández-Vivó, M. (2020). Supervisor feedback: Perceptions from physical education teacher candidates. *Physical Educator, 77*(3), 553-574.

Grudnoff, L. (2011). Rethinking the practicum: Limitations and possibilities. *Asia-Pacific Journal of Teacher Education, 39*(3), 223-234.

Gurl, T. J. (2019). Classroom practices of cooperating teachers and their relationship to collaboration quality and time: Perceptions of student teachers. *Teaching Education, 30*(2), 177-199.

Hanuscin, D. L., Lee, M. H., & Akerson, V. L. (2011). Elementary teachers' pedagogical content knowledge for teaching the nature of science. *Science Education, 95*(1), 145-167.

Hattie, H., & Timperley, J. (2007). The power of feedback. *Review of Educational Research, 77*(1), 81-112.

Hellison, D. (2003). Teaching personal and social responsibility in physical education. In S. Silverman & C. Ennis (Eds.), *Students learning in physical education: Applying research to enhance instruction* (pp. 241-254). Human Kinetics.

Holbrook, K. R. (2022). *Measuring students' perceptions of student teaching university supervisors: Scenario-based scale development using rasch and guttman facet theory* [Doctoral dissertation]. Boston College.

Hudson, P. (2014). Feedback consistencies and inconsistencies: Eight mentors' observations on one preservice teacher's lesson. *European Journal of Teacher Education, 37*(1), 63-73.

Hudson, P. (2016). Forming the mentor-mentee relationship. *Mentoring & Tutoring: Partnership in Learning, 24*(1), 30-43.

Hudson, P., Skamp, K., & Brooks, L. (2005). Development of an instrument: Mentoring for effective primary science teaching. *Science Education, 89*(4), 657–674.

Hume, A., & Coll, R. K. (2007). *The influence of a standards-based qualification on student inquiry in science*. Paper presented at the annual meeting of the National Association for Research in Science Teaching, New Orleans, LA.

Jing-Jing, H. U. (2014). A critical review of pedagogical content knowledge components: Nature, principle and trend. *International Journal of Education and Research, 2*(4), 411-424.

Jones, C. K., Kelsey, K. D., & Brown, N. R. (2014). Climbing the steps toward a successful cooperating teacher/student teacher mentoring relationship. *Journal of Agricultural Education, 55*(2), 33-47.

Kahan, D., Sinclair, C., Saucier, L., & Caiozzi, N. N. (2003). Feedback profiles of cooperating teachers supervising the same student. *Physical Educator, 60*(4), 180-193.

Kastberg, S. E., Lischka, A. E., & Hillman, S. L. (2020). Characterizing mathematics teacher educators' written feedback to prospective teachers. *Journal of Mathematics Teacher Education, 23*(2), 131-152.

Koballa, T. R., & Bradbury, L. U. (2012). Mentoring in support of reform-based science teaching. In B. J. Fraser, K. Tobin, & C. J. McRobbie (Eds.), *Second international handbook of science education* (pp. 361-371). Springer.

Komiskey, C., & Hulse-Killacky, D. (2004). *Supervisors' perceptions about giving and receiving corrective feedback: Implications for counselor education and supervision* [Doctoral dissertation]. University of New Orleans.

MacDougall, L., Mtika, P., Reid, I., & Weir, D. (2013). Enhancing feedback in student-teacher field experience in Scotland: The role of school–university partnership. *Professional Development in Education, 39*(3), 420-437.

Matsko, K. K., Ronfeldt, M., Nolan, H. G., Klugman, J., Reininger, M., & Brockman, S. L. (2020). Cooperating teacher as model and coach: What leads to student teachers' perceptions of preparedness?. *Journal of teacher education, 71*(1), 41-62.

McAlister, A., Lee, D., Ehlert, K., Kajfez, R., Faber, C., & Kennedy, M. (2017). *Qualitative coding: an approach to assess inter-rater reliability* (pp. 5-9). Conference: ASEE Annual Conference & Exposition, Columbus, Ohio.

Ministry of National Education [MoNE]. (2017). *General competencies for teaching profession*. Ministry of National Education Republic of Turkey Directorate General for Teacher Training and Development.

Ministry of National Education [MoNE]. (2018a). *İlköğretim kurumları fen bilimleri dersi öğretim programı [Primary education institutions' science instruction program]*. Talim Terbiye Kurulu Başkanlığı.

Ministry of National Education [MoNE]. (2018b). *Teaching practice guidelines*. Ministry of National Education Republic of Turkey Directorate General for Teacher Training and Development.

Moore, R. (2003). Reexamining the field experiences of preservice teachers. *Journal of Teacher Education, 54*(1), 31-42.

Nguyen, H. T. (2009). An inquiry-based practicum model: What knowledge, practices, and relationships typify empowering teaching and learning experiences for student teachers, cooperating teachers and college supervisors?. *Teaching and Teacher Education, 25*(5), 655-662.

O'Connor, C., & Joffe, H. (2020). Intercoder reliability in qualitative research: debates and practical guidelines. *International Journal of Qualitative Methods, 19*, 1-13.

Paker, T. (2008). Problems of student teachers regarding the feedback of university supervisors and mentors during teaching practice. *Pamukkale Üniversitesi Eğitim Fakültesi Dergisi, 1*(23), 132.

Patton, M. Q. (2014). *Qualitative research & evaluation methods: Integrating theory and practice*. Sage Publications.

Phillippi, J., & Lauderdale, J. (2018). A guide to field notes for qualitative research: Context and conversation. *Qualitative Health Research, 28*(3), 381-388.

Polat, K., Eyceyurt Türk, G., & Altaylı Özgül, D. (2020). Investigation of teaching practice course in school, faculty and preservice teacher dimensions. *Eurasian Journal of Teacher Education, 1*(2), 129-147.

Portelance, L., Caron, J., & Martineau, S. (2016). Collaboration through knowledge sharing between cooperating teachers and university supervisors. *Brock Education: A Journal of Educational Research and Practice, 26* (1), 36-51.

Puttick, S., & Wynn, J. (2020). Constructing 'good teaching' through written lesson observation feedback. *Oxford Review of Education, 47*(2), 152-169.

Ranade, M. (2008). Science education in India. In R. K. Coll & N. Taylor (Eds.), *Science education in context: An international perspective of context on science curricula development and implementation*, (pp. 99-114). Sense.

Range, B., Duncan, H., & Hvidston, D. (2013). How faculty supervise and mentor preservice teachers: Implications for principal supervision of novice teachers. *International Journal of Educational Leadership Preparation, 8*(2), 43-58.

Ritter, J. K., Powell, D., Hawley, T. S., & Blasik, J. (2011). Reifying the ontology of individualism at the expense of democracy: An examination of university supervisors' written feedback to student teachers. *Teacher Education Quarterly, 38*(1), 29-47.

Saka, M. (2019). Evaluations of science teachers regarding the classes of school experience and teaching practices. *Elementary Education Online, 18*(1), 127-148.

Sandvik, L. V., Solhaug, T., Lejonberg, E., Elstad, E., & Christophersen, K. N. (2019). Predictions of school mentors' effort in teacher education programmes. *European Journal of Teacher Education, 42*(5), 574-590.

Scheeler, M. C. (2008). Generalizing effective teaching skills: The missing link in teacher preparation. *Journal of Behavioral Education, 17*(2), 145-159.

Scheeler, M. C., Ruhl, K. L., & McAfee, M. K. (2004). Providing performance feedback to teachers: A review. *Teacher Education and Special Education, 27*(4), 59–70.

Schwartz, C., Walkowiak, T. A., Poling, L., Richardson, K., & Polly, D. (2018). The nature of feedback given to elementary student teachers from university supervisors after observations of mathematics lessons. *Mathematics Teacher Education and Development, 20*(1), 62-85.

Shantz, D. (1995). Teacher education: Teaching innovation or providing an apprenticeship? *Education, 115*(3), 339-344.

Shulman, L. S. (1986). Those who understand: Knowledge growth in teaching. *Educational Researcher, 15*(2), 4-14.

Shulman, L. S. (1987). Knowledge and teaching: Foundations of the new reform. *Harvard Educational Review, 57*(1), 1-23.

Sim, C. (2010). Sustaining productive collaboration between faculties and schools. *Australian Journal of Teacher Education, 35*(1), 518-528.

Smith, K., & Lev-Ari, L. (2005). The place of the practicum in preservice teacher education: The voice of the students. *Asia-Pacific Journal of Teacher Education, 33*(3), 289-302.

Soares, A., & Lock, R. (2007). Preservice science teachers' perceptions of written lesson appraisals: The impact of styles of mentoring. *European Journal of Teacher Education, 30*(1), 75-90.

Steadman, S. C., & Brown, S. D. (2011). Defining the job of university supervisor: A department-wide Study of University Supervisors' Practice. *Issues in Teacher Education, 20*(1), 51-68.

Tang, S. Y. F., & Chow, A. W. K. (2007). Communicating feedback in teaching practice supervision in a learning-oriented field experience assessment framework. *Teaching and Teacher Education, 23*(7), 1066-1085.

Tarekegn, G., Terfa, D., Tadesse, M., Atnafu, M., & Alemu, M. (2020). Ethiopian preservice primary science teachers' perceptions of mentoring in science teaching. *Journal of Science Teacher Education, 31*(8), 894-913.

Tillema, H. H. (2009). Assessment for learning to teach: Appraisal of practice teaching lessons by mentors, supervisors, and student teachers. *Journal of Teacher Education, 60*(2), 155-167.

Topkaya, Y., Tokcan, H., & Kara, C. (2012). Opinions of prospective social studies teachers' views about teacher practice course. *The Journal of Academic Social Science Studies, 5*(7), 663-678.

Van Zee, E. H., Lay, D., & Roberts, D. (2003). Fostering collaborative inquiries by prospective and practicing elementary and middle school teachers. *Science Education, 87*(4), 588-612.

Vertemara, V., & Flushman, T. (2017). Emphasis of university supervisor feedback to teacher candidates. *Journal of Student Research, 6*(2), 45-55.

Weiss, E. M., & Weiss, S. (2001). Doing reflective supervision with student teachers in a professional development school culture. *Reflective practice, 2*(2), 125-154.

Wilson, E. K. (2006). The impact of an alternative model of student teacher supervision: Views of the participants. *Teaching and Teacher Education, 22*(1), 22-31.

Won, N., Liu, K., & Bukko, D. (2019). Developing instructional skills: Perspectives of feedback in student teaching. *Networks: An Online Journal for Teacher Research, 21*(2), 8.

Yayli, D. (2008). Theory-practice dichotomy in inquiry: Meanings and preservice teacher mentor teacher tension in Turkish literacy classrooms. *Teaching and Teacher Education, 24*(4), 889-900.

Yin, R. K. (2003). *Case study research: design and methods* (3rd ed.) Sage Publications.

Journal of Theoretical Educational Science, 17(1), 169-201, January 2024
Kuramsal Eğitimbilim Dergisi, 17(1), 169-201, Ocak 2024
[Online]: http://dergipark.org.tr/akukeg
DOI number: http://doi.org/10.30831/akukeg.1312055

Interdisciplinary Gender Equality Education Integrating Science, Mathematics and Information Technologies and Software Courses: A Sample from Türkiye*

Matematik, Fen Bilgisi ve Bilişim Teknolojileri ve Yazılım Dersleri ile Bütünleştirilmiş Disiplinlerarası Toplumsal Cinsiyet Eşitliği Eğitimi: Türkiye'den Bir Örnek

Buket TURHAN TÜRKKAN** **Nihan ARSLAN NAMLI*****
Betül KARADUMAN**** **Memet KARAKUŞ*******

Received: 09 June 2023 **Research Article** **Accepted:** 04 January 2024

ABSTRACT: The general aim of the study was to improve students' awareness of gender equality. An instructional design was developed in which mathematics, science, information technologies and software (ITS) courses were integrated, with the theme of gender equality at the center. The method of the study was determined as a case study. The participants of the study were 21 sixth-grade students at a public school in Turkey. Research data were collected using the "Views Form on Gender Equality" and "Practice Evaluation Form" developed by the researchers. According to the students' views, it was determined that an interdisciplinary education on gender equality had positive effects on improving students' awareness of gender equality. When the students' views about the practice process were examined, it was seen that the positive opinions about the process were high in number, and students presented few negative views.

Keywords: Gender equality, mathematics course, science course, information technologies and software course, sixth-grade students, interdisciplinary approach.

ÖZ: Çalışmanın genel amacı, öğrencilerin toplumsal cinsiyet eşitliği konusunda farkındalıklarını artırmaktır. Matematik, fen bilgisi ve bilişim teknolojileri ile yazılım derslerinin bütünleştirildiği, toplumsal cinsiyet eşitliği temasının merkeze alındığı bir öğretim tasarımı geliştirilmiştir. Araştırmanın yöntemi durum çalışması olarak belirlenmiştir. Araştırmanın katılımcıları Türkiye'de bir devlet okulunda öğrenim gören 21 altıncı sınıf öğrencisidir. Araştırma verileri, araştırmacılar tarafından geliştirilen "Toplumsal Cinsiyet Eşitliğine İlişkin Görüş Formu" ve "Uygulama Değerlendirme Formu" ile toplanmıştır. Öğrenci görüşlerine göre toplumsal cinsiyet eşitliğine yönelik disiplinlerarası bir eğitimin öğrencilerin toplumsal cinsiyet eşitliğine ilişkin farkındalıklarını geliştirmede olumlu etkileri olduğu belirlenmiştir. Öğrencilerin uygulama sürecine ilişkin görüşleri incelendiğinde, sürece ilişkin olumlu görüşlerin fazla olduğu ve öğrencilerin az sayıda olumsuz görüş bildirdiği görülmüştür.

Anahtar kelimeler: Toplumsal cinsiyet eşitliği, matematik dersi, fen bilgisi dersi, bilişim teknolojileri ve yazılım dersi, altıncı sınıf öğrencileri, disiplinlerarası yaklaşım.

* A part of this study was presented at the 7th International Congress on Curriculum and Instruction (October 9-12, 2019, Ankara, Türkiye).

** *Corresponding Author*: Asst. Prof. Dr., Çukurova University, Adana, Türkiye, bturhan@cu.edu.tr, https://orcid.org/0000-0003-2528-4022
*** Lect. Dr., İskenderun Technical University, Hatay, Türkiye, nihan.arslannamli@iste.edu.tr, https://orcid.org/0000-0002-5425-1468
**** Res. Asst. Dr., Çukurova University, Adana, Türkiye, bkaraduman@cu.edu.tr, https://orcid.org/0000-0001-7724-3930
***** Assoc. Prof. Dr., Çukurova University, Adana, Türkiye, memkar@cu.edu.tr, https://orcid.org/0000-0002-6099-5420

Citation Information
Turhan Türkkan, B., Arslan Namlı, N., Karaduman, B., & Karakuş, M. (2024). Interdisciplinary gender equality education integrating science, mathematics and information technologies and software courses: A sample from Türkiye. *Kuramsal Eğitimbilim Dergisi [Journal of Theoretical Educational Science], 17*(1), 169-201.

ISSN: 1308-1659

It has been indicated that inequalities increase at both national and international levels due to the differentiation of societies on the basis of qualities such as race, class, gender, and language (Kellogg, 2002). One of these areas of inequality is gender inequality. The concept of sexism is defined as the belief that women and men have different characteristics and that one gender has the right to more power and resources than the other, and it is stated that the attitudes and behaviors created by sexism in children have negative effects on both genders (Schniedewind & Davidson, 2006). In addition, it is emphasized that individuals are also adversely affected by school culture in terms of gender roles (Levine et al., 1995). In this context, gender-based prejudices that some students have negatively affect the other students at school as well. Besides, it is put forward that individuals could make gender discrimination even without noticing while they are avoiding sexist behaviors (Schniedewind & Davidson, 2006). Moreover, it is determined that there are also problems regarding gender inequality in Turkey (Gözütok et al., 2017).

There could be an interpretation regarding the present situation of gender inequality in Turkey when the research and statistical data on this matter were examined. In the Global Gender Gap Report, Turkey is at 130th rank among 153 countries in terms of the gender gap index according to the 2020 ranking (World Economic Forum [WEF], 2019). On the other hand, in the Global Gender Gap Report 2021, Turkey is 133rd among 156 countries (WEF, 2021). When the net schooling rates in higher education in Turkey were analyzed, this rate was 40,6% for men and 46,3% for women (Turkish Statistical Institute [TURKSTAT], 2021a). According to the results of the household labor force survey, in 2019, the employment rate in Turkey was 28.7% for women and 63.1% for men (TURKSTAT, 2021c). When the data for R&D employment were examined, 68,1% of sectoral distribution by gender is men while 31,9% is women (TURKSTAT, 2021b). Turkey is behind EU countries in terms of girls' gross enrollment rates and gender equality index at all levels of education and expected years of education for girls from primary to tertiary education (Maya, 2013). Based on these results, it could be said that there are inequalities regarding gender in areas such as education, employment, and labor in Turkey. These data show the problems related to gender equality, yet producing solutions for them and applying them are also important subjects, in addition to determining the problems. It has been indicated that it is necessary to question the role of education in struggling with social problems and the formation of a more social society reflecting the values regarding personal freedoms (Apple, 2017). In order to solve social problems, the social function of education should be activated (Turhan Türkkan, 2017). In this sense, various studies investigating the perceptions of students and teachers regarding gender equality and examining the effects of curricula for ensuring gender equality were conducted.

In a study at the primary level, it was determined that 7-8-year students have gender stereotypes in toy choice (Bağçeli Kahraman & Başal, 2011). It was put forward in a study on fourth graders that students possess traditional stereotypes and views on women being weaker and more powerless (Yolcu, 2021). In another study with fourth graders, it was also found that students have gender stereotypes towards occupations and households (Yolcu & Sarı, 2018). In a study conducted on the gender perceptions of primary school students between the ages of 7-15, it was determined that the gender perception of children changes according to the age period; however, being a man is

seen as more valuable in every period (Kılıç et al., 2014). Through a study at the middle school level, it was revealed that middle school students do not know gender concepts, and they adopted traditional gender roles (Yeşil & Balcı Karaboğa, 2021). Within another study at the same level, it was put forward that eight graders have views on gender inequality (Kalaycı & Hayırsever, 2014). Besides these studies, during the needs analysis of this study, similar results were found, and it was determined that the awareness of sixth graders towards gender equality is not sufficient, and there is a need to develop this awareness (Turhan Türkkan et al., 2018a, 2018b). In a study conducted at the high school level, it was found that high school students have educational needs on the concept of gender, gender roles and stereotypes, homosexuality, women's participation in decision-making mechanisms, violence against women, and women's participation in working life and income (Acar-Erdol & Gözütok, 2017). Additionally, studies have presented evidence that pre-service teachers hold gender stereotypes and views on gender inequality (Acar-Erdol et al., 2019; Aslan, 2015; Koyuncu Şahin et al., 2018). Moreover, it was seen that similar results were reached through studies with bachelor students as well (Öngen & Aytaç, 2013; Sis-Çelik et al., 2013; Vefikuluçay et al., 2007; Vefikuluçay Yılmaz et al., 2009). Based on these studies in Turkey, it is considered that many students from primary to higher education have stereotypes and negative views on gender equality. In this respect, it can be said that there is a need for a national education on gender equality.

Within the studies at the primary and middle school level, it has been concluded that gender equality education is effective in decreasing students' prejudices towards gender, developing concept knowledge about gender, changing perspectives on gender roles, and raising awareness of gender equality (Akita & Mori, 2022; Brinkman, 2009; Seçgin & Kurnaz, 2015; Yeşil & Balcı Karaboğa, 2021; Yolcu, 2021). Similar results were reached through also studies related to gender equality education with pre-service teachers (Acar-Erdol, 2019; Aydemir, 2019; Esen, 2013; Şener Özel, 2019). Another study conducted with teachers determined that teachers' knowledge about gender differences of students who attended training, their perceptions of justice towards the diversity in the classroom, and their knowledge skills to encourage girls and boys increased (Kollmayer et al., 2020). Besides, it was found that the women's studies course held at the higher education level positively affected the students' views on gender equality (Harris et al., 1999; Stake, 2006; Stake & Hoffmann, 2001; Thomsen et al., 1995). Based on the results of these studies, it is thought that the applied practices have positive effects in terms of gender equality. The studies conducted were carried out mostly with adults. In addition to this, the practices were either developed directly on the subject of gender equality or were associated with courses in social fields such as social studies and life sciences.

When the research conducted within the Turkish educational system was analyzed, it was put forward that the traditional female role approach in society also exists in the education field. Teachers and administrators reinforce this with the language used in-class processes, activities, and course materials - mostly without noticing-and this situation also interrupts women's education in particular, contributing to women's subsidiary position (Polat, 2010). In the studies carried out in this context, it is stated that there are deficiencies regarding gender equality in textbooks and curricula (Çelik et al., 2019; Güney, 2016; Işık Demirhan, 2021; Kalaycı & Hayırsever, 2014;

Karakuş et al., 2018; Kükrer & Kıbrıs, 2017). With respect to this, it becomes clear that there is a need for arrangements regarding gender equality.

In order for individuals to live in an equal world, it is thought that it will be beneficial both individually and socially for individuals to receive gender equality education to produce creative and applicable solutions to these problems and realize the problems related to gender equality. Besides this, this education is mostly given in courses within the social field. There are very few studies handled in numerical courses such as mathematics, science, and ITS. In this sense, the social integration of mathematics, science, and ITS courses will be a unique practice both in terms of practice and in terms of literature.

Mathematics course is considered a tool for dealing with social issues in education (Gutstein, 2007a). Equality and social justice can be taught in mathematics classes to help students interpret and apply mathematical knowledge to answer questions that will potentially strengthen their lives and the society in which they live (Leonard et al., 2010). For this purpose, the events in the world could be approached from a mathematical point of view (Noyes, 2007), and besides caring for numbers, they could be studied with numerical data regarding social topics such as unemployment, racism, and inequality (Miner, 1995). In this context, it is thought that, in general, social topics, in particular gender equality topics, could be handled in mathematics course. Within this scope, when the studies integrating mathematics with social topics were examined, it was determined that there conducted some practices for middle school level (Allen, 2003; Gutstein, 2003; Turhan Türkkan, 2017), for high school level (Brantlinger, 2007; Harper, 2017; McNamee, 2013; Voss, 2015; Wonnacott, 2011), for pre-service teachers (Johnson, 2005; Koestler, 2010) and for teachers (Gonzalez, 2009; Wager, 2008) and these studies have constructed positive effects on individuals regarding social justice and equality. Based on this, it is considered that including topics such as equality, justice, social justice, and gender equality in mathematics course would provide positive outcomes for students.

The main purpose of science courses is to raise science-literate individuals. One of the important components of science literacy is the ability to establish the relationship between science and technology, society, and the environment. At this point, it is stated that science courses are taught, putting social issues at the center. Students get an active role and use their scientific knowledge in problem-solving, so they will make progress regarding the nature of science (Pedretti, 1999). As for Tal and Kedmi (2006), when students learn about social issues about science and their reflections on society, they can analyze the incidents, make decisions, and develop attitudes towards this subject. It is indicated that significant research centers such as the National Research Council and Queensland School Curriculum Council emphasize the necessity of enhancing the skills of students regarding their ability to discuss, analyze socio-scientific matters, and make decisions based on knowledge (Topçu et al., 2014). It is thought that teachers who incorporate social justice into their lessons will ideally create a classroom environment that helps students be effective voices for change while supporting students to question the unfair power relations in society (Gutstein, 2007b; Mayberry, 1998; Roth, 2007). At this point, it is considered that including social topics in science course curricula and increasing their awareness in this matter positively effect students' being science literate, which is the main purpose of science courses. History of

science has an important place in teaching scientific literacy (Tokuş, 2018). An individual who is aware of historical developments in science becomes aware of how discoveries and inventions were made, what happened in this process, what conditions were effective, and the efforts made in this way, and in this respect, the history of science is considered a very suitable subject to be used in a science course (Laçin Şimşek, 2009). It can be said that covering the social context in terms of the history of science in the science course is also important in terms of integrating science and social issues.

Rapid changes in technology indisputably bring up concepts such as social justice, equality, and gender equality. There is a technology gap called the digital gap between communities or generations where there are inequalities in access to information technology (Tarman et al., 2015). In situations with a technology gap, since students cannot access information equally, it becomes impossible to expect the same results from students. Hellsten (2007) concluded that the aim of technology is actually to make students one; however, it creates a digital divide these days. The technology could be used to provide social justice education and expand opportunities in this sense, and it is recommended that research be conducted concerning the use of hidden social justice curricula in traditional educational settings (Mitchell, 2015). Hence, it is thought that including concepts such as equality, social justice, and gender equality in ITS course would provide significant contributions to the literature. In addition, one of the aims of ITS curriculum is to *"Develop innovative and original projects for the solution of problems encountered in daily life (problems faced by elderly and disabled individuals, etc.)"* (Ministry of National Education, 2018) and it is thought that ITS curriculum includes the integration of social issues with the subjects in the course.

Considering the multidimensionality of real-life problems such as equality, justice, and social justice, handling these matters in an interdisciplinary approach is crucial in terms of reaching solutions for the problems. When the philosophy and approaches underlying contemporary curricula were counted, it was considered necessary to deal with the subjects within the framework of an interdisciplinary approach and to make connections between courses on real-life subjects (Karakuş et al., 2017). In the interdisciplinary approach, various courses in the curriculum are brought together, and the main point in this approach is that the curriculum developers try to use a series of course-based perspectives (Jacobs, 1989). The interdisciplinary approach is considered the integration and interaction of the issues involved in solving a common problem and involves synthesizing and integrating interrelated disciplines into a coordinated and coherent whole (Munkebye et al., 2020). In an interdisciplinary curriculum, connections between disciplines are strong and evident; issues such as 21st-century skills, global problems, sustainability, and intercultural skills can be addressed with an interdisciplinary approach, and in such a case, the boundaries between disciplines become blurred (Drake & Reid, 2020). Regarding climate change, inequalities in different areas, and social problems that concern the whole world and are among the sustainable development goals, it is necessary to study together across disciplines, develop an interdisciplinary understanding, and even produce solutions to problems with understandings beyond disciplines (İnci & Kaya, 2022). It is stated that there is a need for an interdisciplinary curriculum design that takes science, mathematics, and information technology courses together and reflects the real world

(Millar, 2020). In this respect, it is thought that teaching gender equality and establishing interdisciplinary connections with mathematics, science, and ITS courses will make a contribution. Besides, the integration of these three courses with gender equality topics would give it a different perspective. Moreover, no practice or study towards gender equality integrating mathematics, science, and ITS has been found. In all these respects, it is thought that this study will fill the gap in the literature and make important contributions to the field with new information. Within this context, the main purpose of this study was to develop, apply, and evaluate an instructional design for gender equality education. Based on this main purpose, the study aimed to develop sixth graders' awareness on gender equality. In this sense, answers to the questions below were sought:

1. How does an interdisciplinary teaching on gender equality in the sixth grade of middle school contribute to students' awareness of gender equality?
2. What are the students' views on gender equality education at the sixth-grade level of middle school?

Method

In this section, the dimensions of the research model, participants, application process, data collection, and data analysis are included.

Research Design

The research was conducted through the case study method. A case study is a detailed investigation of an incident or several incidents with any convenient method. The main aim of a case study is to understand every aspect of an incident (Punch, 2005). Case studies could be conducted via quantitative or qualitative approaches. Through qualitative case studies, the factors regarding a case are investigated in a wholistic way and focus on how these factors affect the case in question and how these factors are affected by the case. In addition to this, when the changes and the processes of the case are seen as important, it might be possible to study the cases in the long term (Yıldırım & Şimşek, 2008). In this research, the application of the developed design is considered as a case. In this context, the design development process has been taken into account.

Participants

The participants of the study consisted of 21 students studying in sixth grade through the second semester of the 2018-2019 education year in a state school within a middle-socioeconomic region of Adana province in Turkey. Of the 21 sixth-grade students who made up the participants, 14 were girls, and 7 were boys.

The school where the research was conducted is located in Sarıçam district, in the east of Adana. There are approximately 60 teachers and 800 students at the school. There are 20 classrooms in the school. The region where the school is located is a newly formed region in Adana, and the people living in this region are generally at a middle socio-economic level.

Social issues can be included in teaching processes starting from the elementary school level (McBee, 1996). Dealing with social issues is considered important at the secondary school level (Schniedewind & Davidson, 2006). Students at the elementary school level can understand social issues, which can be included in the period between

the ages of 12-18 (Peterson, 2002). Considering that the issue of gender equality also has a social context, it was thought that a study based on this issue could be conducted at the secondary school level, and therefore, the research was conducted at the sixth-grade level.

Instruction Process

At the beginning of this design, a needs analysis was first made. In line with the needs analysis results, teaching attainments were determined first. Then, these attainments were organized within the framework of relevant disciplines and subjects, and content was formed. In the next stage, the variables, teaching methods, and learning and teaching process activities were arranged. In the last stage, arrangements were made for how teaching would be evaluated. After the theoretical part of the design process was completed, practice started.

The process of instruction lasted for 8 lesson hours in total. The instruction was conducted in April and May of the 2018-2019 education year. In the first instructional process, basic information and concepts were handled within the "Introduction to Gender Equality Topic." Through the second instructional process, the topic "Education and Gender Equality" was taught. The topics included in this course were integrated with the mathematics course. An example of integrating gender equality into the mathematics course can be presented as follows: The "Solves and poses problems that require performing four operations with natural numbers" outcome in the mathematics curriculum is rearranged as "Solves mathematical problems regarding gender discrimination in education" in this practice. In this example, the issue of gender equality in education is integrated with the topic of solving problems with natural numbers in mathematics. In the third instructional process, the topic of "history of science and gender equality" was dealt with and integrated with the history of science topic within the science course. As for the fourth instructional process, "Art, science, literature, social media, and gender equality" was the topic. The topics taught in this course were related to information technologies and software. The activities and used materials are shown in Table 1.

Table 1

The Activities and Used Materials in Instruction

The Activities and Used Materials in Instruction	
Introduction to Gender Equality Topic	• 4 leaf clover event • Video presentations regarding gender equality. • Presentation of the concepts regarding sex, gender, gender equality, gender discrimination, gender stereotypes • Worksheet 1 (Basic concepts) • Evaluation Sheet 1
Education and Gender Equality	• Video about equal opportunity in education • Discussion on the barriers girls face in education • Worksheet 2 (Primary and secondary enrollment rates for various countries – gender equality in mathematics and education) • Worksheet 3 (Solving math problems that involve calculating natural numbers with gender equality situations) • Video presentation and discussion about Malala's life • Evaluation Sheet 2
History of Science and Gender Equality	• Discussion on the concept of "scientist" • Cartoon and discussion about the life of Marie Curie • Video and discussion on the life of Prof. Dr. Dilhan Eryurt • Worksheet 3 (Occupations and gender equality) • Evaluation Sheet 3
Sports, Arts, Literature, Social Media and Gender Equality	• Digital Story Presentation (Stories inspired by the lives of Geeta Phogat and Mary Shelley) • Examining social media posts in terms of gender equality • Evaluation with Kahoot! application (Social media posts involving gender equality/discrimination)

Data Collection and Analysis

The data of the study were collected via the "Opinion Form on Gender Equality" and "Instruction Evaluation Form" developed by the researchers. Both forms were examined by experts studying gender equality. Students put their views on both forms in written form.

Opinion Form on Gender Equality consists of five questions developed by the researchers towards the needs analysis of the study (Turhan Türkkan et al., 2018a). An example case, including gender equality/inequality, was presented in each question. The first case is about gender inequality regarding occupations. The second case is about gender equality in households, the third one is about gender equality in sports, the fourth case is about gender equality in promotion, and the fifth is about gender equality in terms of the right to education. While there is an evident statement of gender discrimination in the first and fourth cases, there aren't any statements of gender discrimination in the second, third, and fifth cases; instead, a hidden situation is present. After each case, a three-choice sub-question was asked of the students. In this direction, the students first marked that they found the situation in the case correct; they were

undecided or did not find it correct. After they stated these opinions, they were asked to express their reasons. Besides asking how they would feel or behave, putting themselves in some individuals' place, they were asked to give the reasons for their views.

As for the Instruction Evaluation Form, there are six questions and three probe questions. These questions were arranged within the scope of their evaluations for the application and were prepared to determine how they found the application, its differences with other courses, the changes in their perspectives on gender equality, their suggestions, and additional opinions.

The inductive analysis method included in content analysis was used to analyze both forms. In this sense, first of all, students' views were transferred to computer. Afterwards, students' views were coded. Themes and sub-themes were formed, bringing the related codes together.

Roles of the Researchers

Since the research was carried out within a project, experts in each context of the research worked together. Since this research included the disciplines of mathematics, science, and ITS, an expert from each field took part in it. A mathematics education expert was actively involved in the planning for mathematics subjects, a science education expert was actively involved in the planning for science subjects, and a computer and instructional technologies education expert was actively involved in the planning for ITS education. In order to ensure and control interdisciplinary connections, a curriculum development expert who studies interdisciplinary curriculum development was assigned. In the gender equality dimension of the research, the project manager, who is also a mathematics education and curriculum development expert, actively studied, and this researcher also has scientific research on the teaching of social issues. The mathematics education expert, who is also the first author of the research and the coordinator of this project, also served as a practicing teacher.

Credibility of the Research

Credibility can be achieved by describing in detail the experiences and procedures regarding the process as a researcher (Cope, 2014). In terms of the credibility of the research, it is stated that strategies such as peer examination, long-term interaction, detailed description and reporting, and controlling researcher effects can be used (Patton, 2002). In this respect, the roles of the researchers were explained in detail, detailed information was provided about the process of conducting the research, an effort was made to ensure that the researchers acted objectively in the process, and a long-term interaction was tried to be established with the participants during the implementation process. Besides, peer examination was applied for the reliability of data analysis. With respect to this, the analyses one of the researchers conducted were examined by the other researchers, and the analyses were completed.

Ethical Procedures

Official permission was obtained from the Adana Provincial Directorate of National Education to conduct the research (25/04/2018, E.25867). In addition, written informed consent was obtained from the parents of the participant students.

Findings

The findings of the study were presented under three headings: findings regarding the awareness status of students for gender equality before the instruction, findings regarding the awareness of students for gender equality after the instruction, and findings regarding students' views on the instruction.

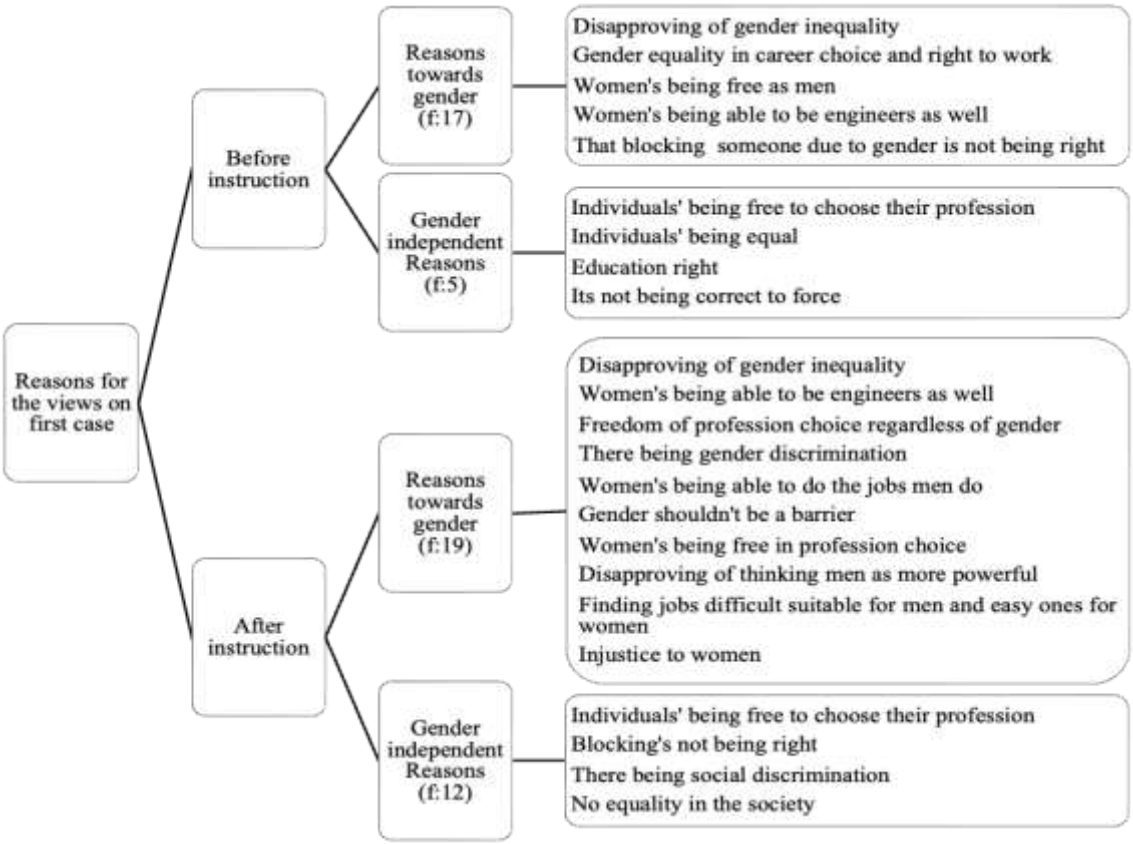
Findings Regarding the Awareness of Students for Gender Equality

Students stated that they did not find the situation correct in the first case, both before and after the instruction. The reasons for students' views on the first case before and after the instruction are presented in Figure 1.

As seen in Figure 1, students presented the reasons for their views on the first case, mostly towards gender, before the instruction. Besides, some students put forward gender-independent reasons as well. Based on this, it could be said that most of the students noticed the gender inequality within the example case in the first question; on the other hand, some students did not notice the inequality. On the other hand, most of the reasons that students presented for their views were related to gender after the instruction. In addition, there were also students who put forward gender-independent reasons. When compared to the situation before the instruction, it was seen that the reasons given by the students for gender increased in number. Based on this, it can be said that the students' awareness of gender enhanced while students questioned the reason for the situation.

Figure 1

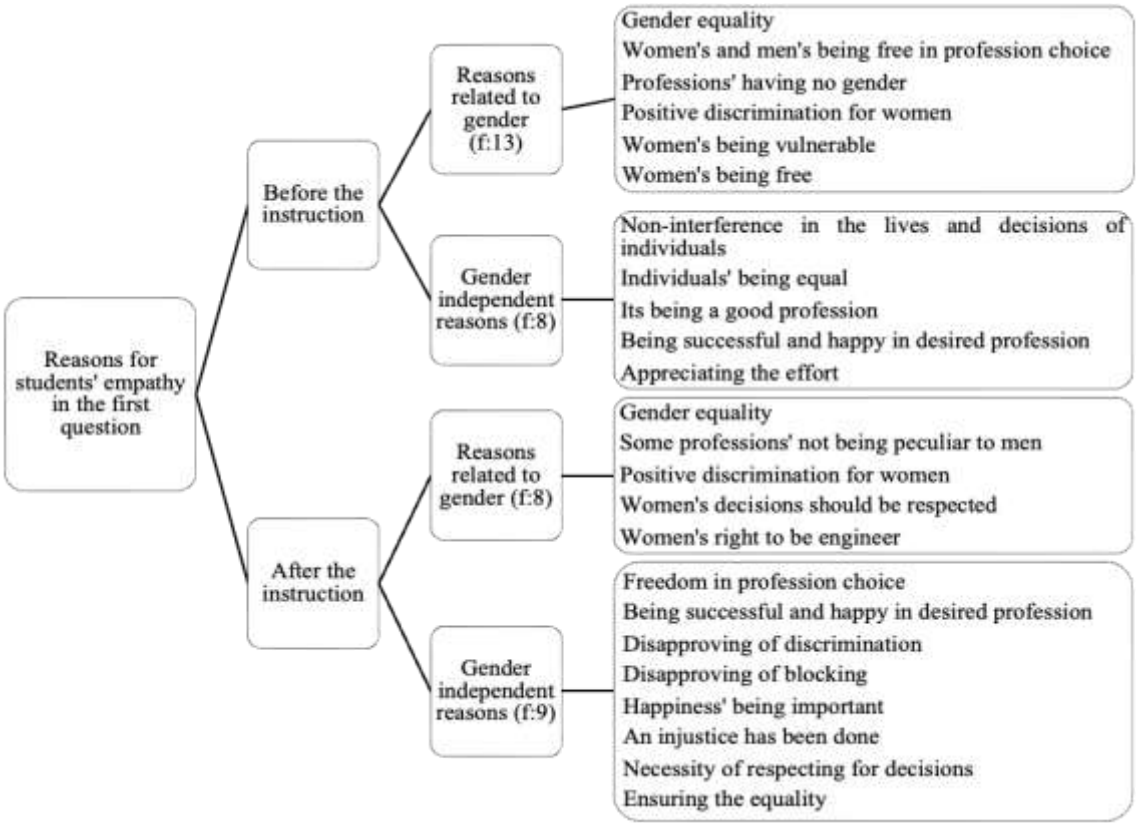
Reasons for Students' Views on the First Case



Before the instruction, in the case of the situation where they put themselves in the place of family, ten students stated that they would allow the girl, six students stated they would support both two children, four students indicated that they should be freed or not intervened and one student told that individual should be asked for his/her wish. On the other hand, after the instruction, in the case when they put themselves in place of family, seven students mentioned that they would let the girl, nine stated that they would support both two children, four students stated they should be freed or not intervened, and one student stated s/he would support girl more. Students' reasons regarding the empathy they showed within the first question were presented in Figure 2.

Figure 2

Reasons for Students' Empathy in the First Question

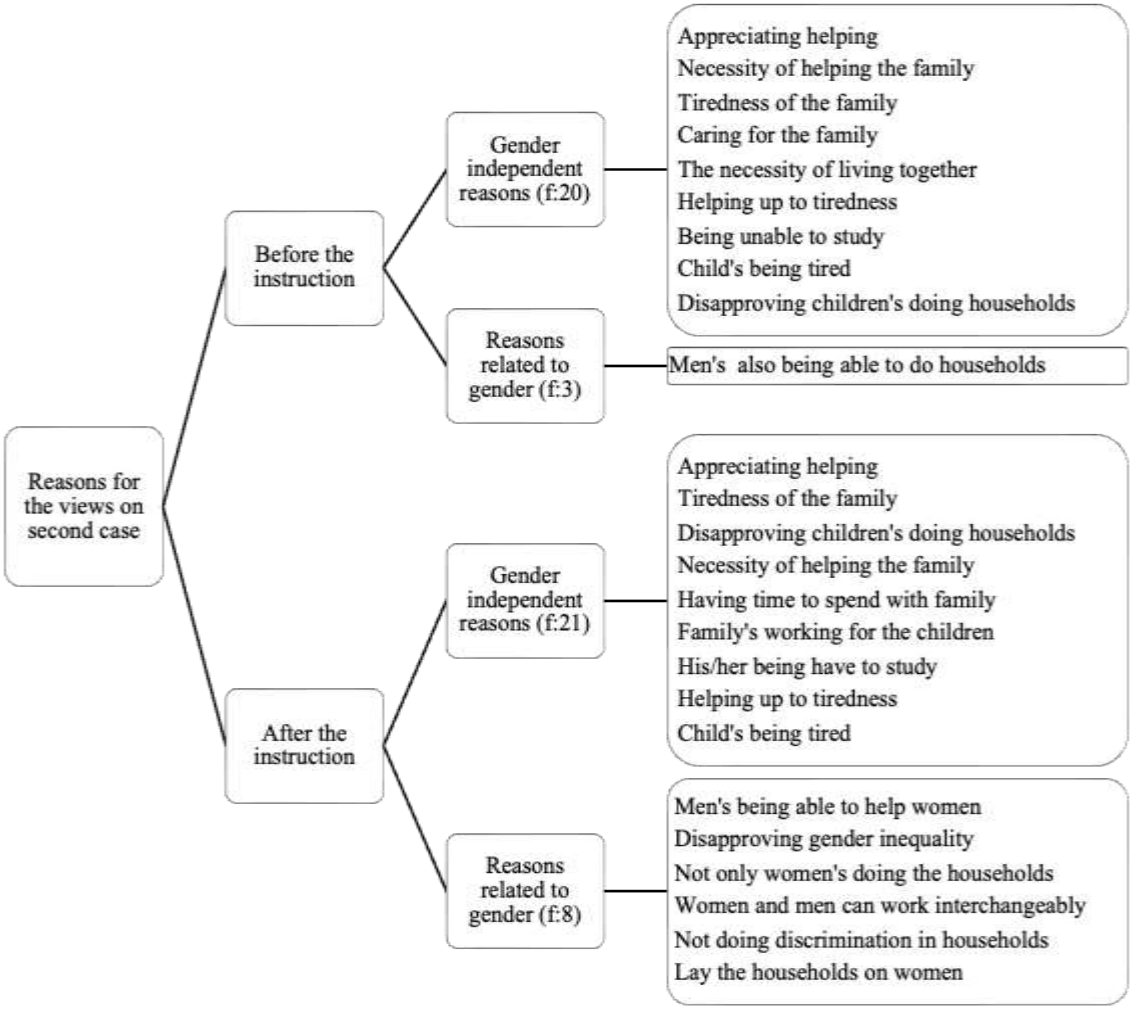


As in Figure 2, the reasons students put forward before the instruction were much more related to gender. Besides, gender-independent reasons were presented as well. In the first question, it was seen that they generally realized the gender inequality in the case. Moreover, it was determined that some students' reasons and rationales were unrelated to gender topics. In this sense, it could be said that some students do not realize gender inequality. After the instruction, while some students presented reasons regarding gender, some of them gave gender-independent reasons.

For the case in the second question, before the instruction, 18 students mentioned they found the behavior right, two students stated that they were undecided, and one student did not find it right. After the instruction, 16 students stated that they found the case in the second question right, three stated they were undecided, and two stated they did not find it right. The reasons for their views are shown in Figure 3.

Figure 3

The Reasons for Students' Views on the Second Question

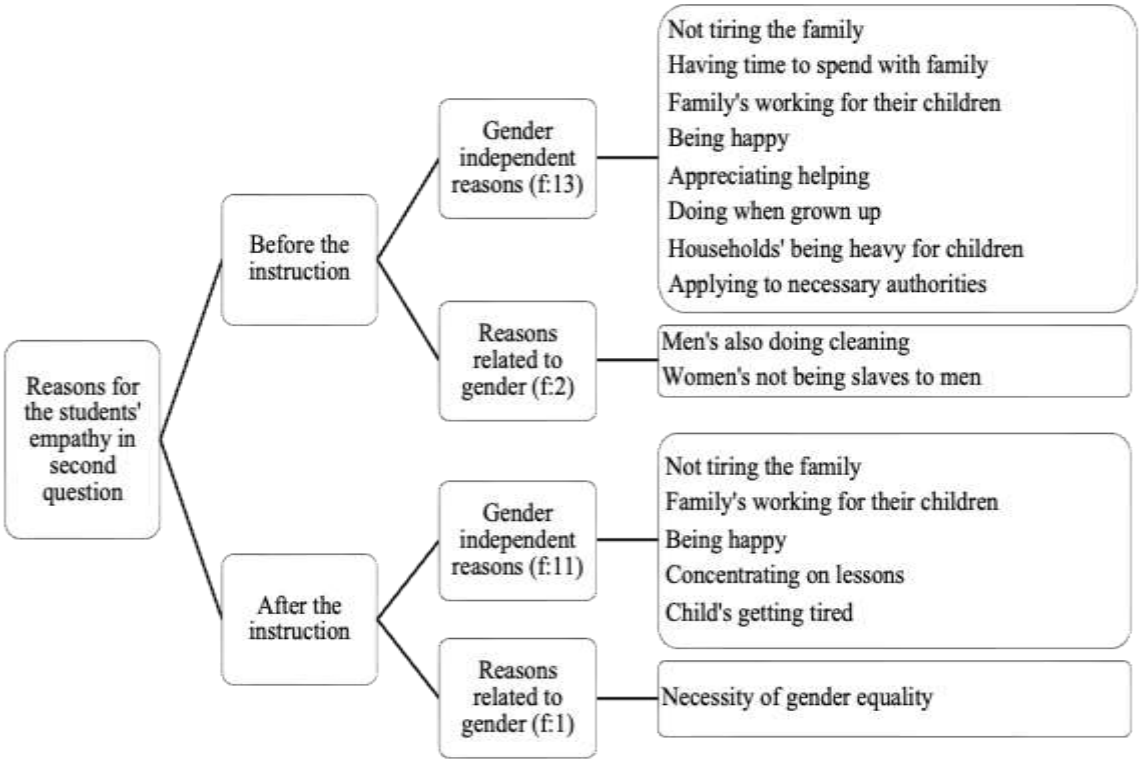


It is seen in Figure 3 that students explained their reasons for views on the second question much more as gender independent before the instruction. Besides, there were few students who came up with reasons related to gender. Based on this, it could be asserted that most of the students did not realize the gender inequality in the case within the second question, while few students realized it before the instruction. After the instruction, most of the reasons students put forward were not related to gender. However, some students presented reasons related to gender as well. Compared to the situation before the instruction, it was seen that the reasons related to gender students mentioned increased in number. It could be said that the students' awareness of gender enhanced while questioning the reasons for the situation.

Before the instruction, 17 students stated that they behaved the same way when they put themselves in Ozan's place; two indicated that they would help with some households, and two said they would not. On the other hand, after the instruction in the situation, they put themselves in Ozan's place; 15 students stated they would behave in the same way, while three students said they would help with some households and three students would not. The reasons students have regarding these views are presented in Figure 4.

Figure 4

Reasons for Empathy of Students within the Second Question

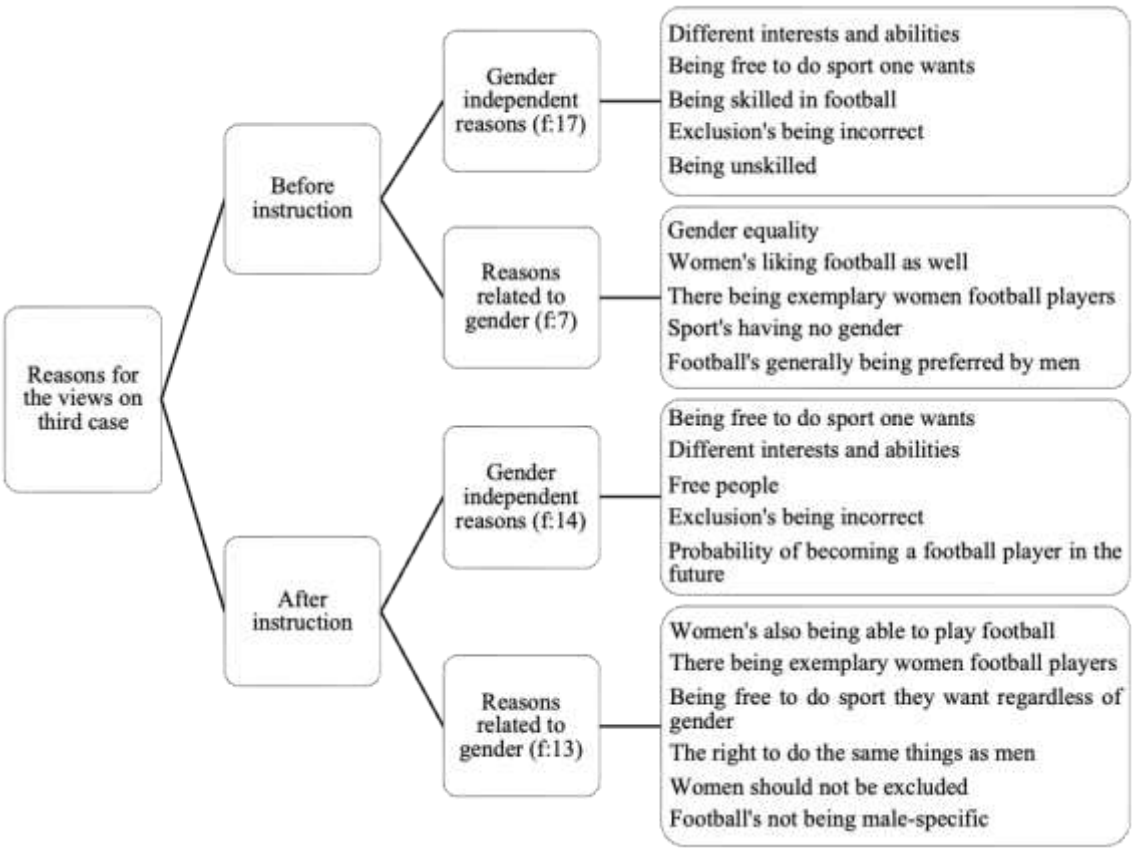


As is seen in Figure 4, most of the reasons students put forward before the instruction were related to gender-independent issues. However, few students explained their reasons as gender-related. In the second question, in general, it was found that most of the reasons and rationales students presented for the situation in the case were not related to gender topic. Based on this, it could be asserted that they did not handle the situation in the case with the gender topic. As for the reasons students brought about after the instruction, most of them were gender independent. However, one student presented a gender-related reason.

Before the instruction, 20 students for the third question stated that they did not find the situation in the sample case correct, while one student expressed that s/he was undecided. After the instruction, all of the students stated that they did not find it correct. The reasons for this opinion are given in Figure 5.

Figure 5

Reasons for the Views on the Third Case

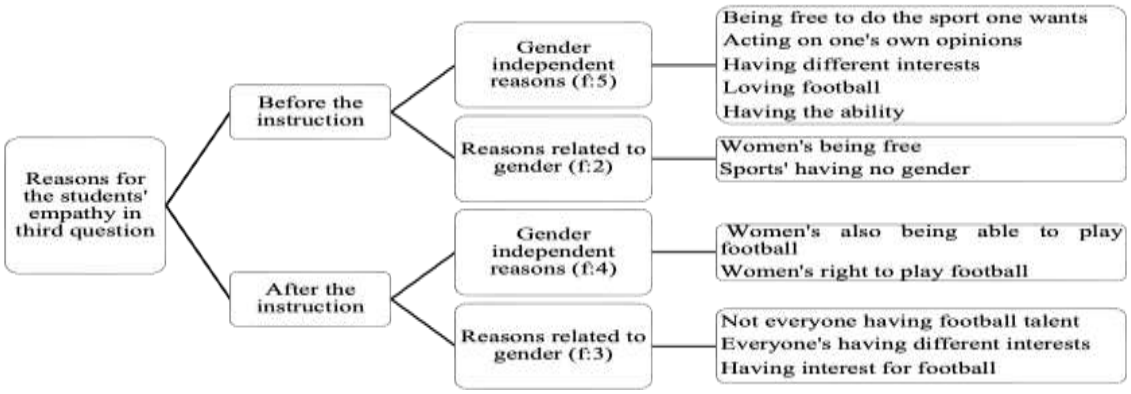


As seen in Figure 5, before the instruction, the students explained the reasons for their opinions on the third question, which was mostly gender independent. Besides, there were also students who related their views to gender topics. Based on this, it can be said that most of the students did not notice the gender inequality in the case of the third question before the instruction, while some students noticed the gender inequality. After the instruction, on the other hand, some of the reasons given by the students were associated with the gender issue, while some were not. Compared to the situation before the instruction, it is seen that the reasons by students related to gender increased in number. In line with this, it could be inferred that their gender equality awareness enhanced while students questioned the reasons for the situation.

Before the instruction, in the case they put themselves in Damla's shoes, eight students expressed that they would play together, six students would support, four students would not be excluded, two students would congratulate, one student would defend, one would want him to teach how to play football, one would not criticize, one would let him alone, one would inform the teacher, one would show empathy and respect. On the other hand, after the instruction, when they put themselves in Damla's place in the case, five students indicated they would respect, four students would not exclude, three students would support, three students would play together, two students would not tease, one student would congratulate, one student would ask her/him to teach him/her, one student would not interfere, and one student would not criticize. The reasons they put forward are presented in Figure 6.

Figure 6

Reasons for the Students' Empathy in the Third Question

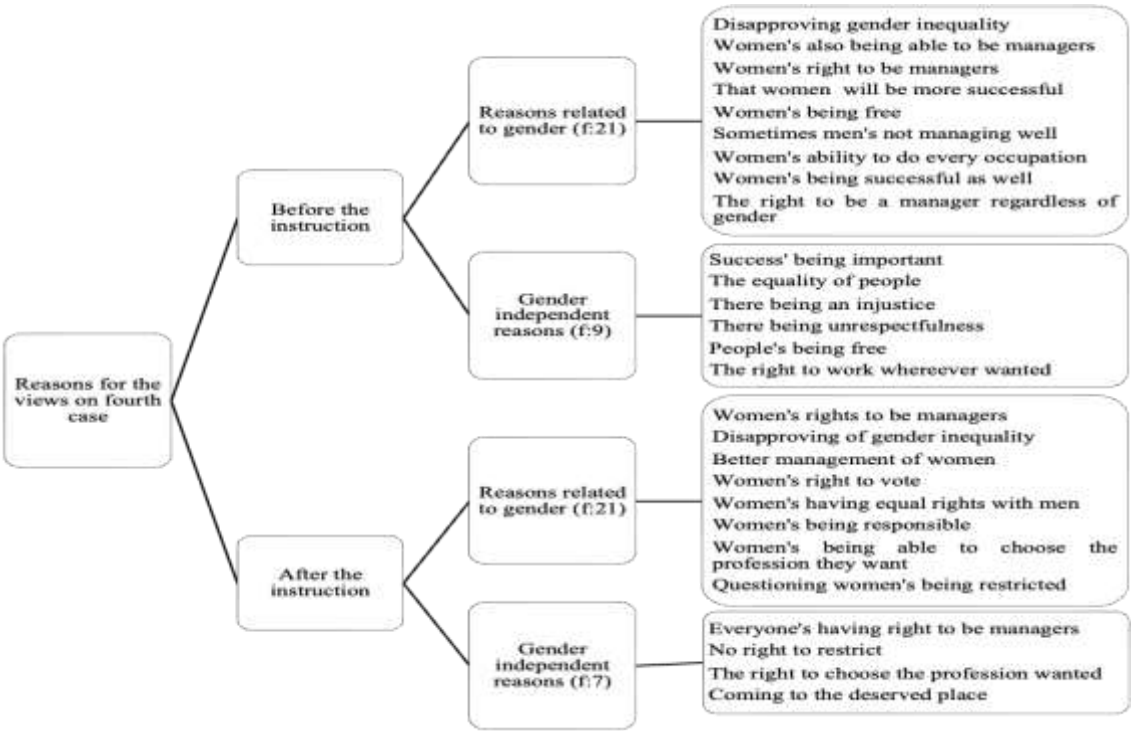


As seen in Figure 6, the reasons given by the students before the instruction were mostly about gender-independent issues. However, there were also students who associated their reasons with gender issues. It could be asserted that students generally did not realize the gender inequality in the case within the third question. In addition to this, it was determined that the reasons and rationales some students came up with related to gender topics. With this respect, it could be said that students did not realize the gender inequality was in the majority. Yet, after the instruction, some of the reasons students were gender independent while some were related to gender.

Before and after the instruction for the fourth question, all of the students stated that they did not find the situation in the case correct. They presented the reasons for their views as in Figure 7.

Figure 7

Reasons for the Views on the Fourth Case

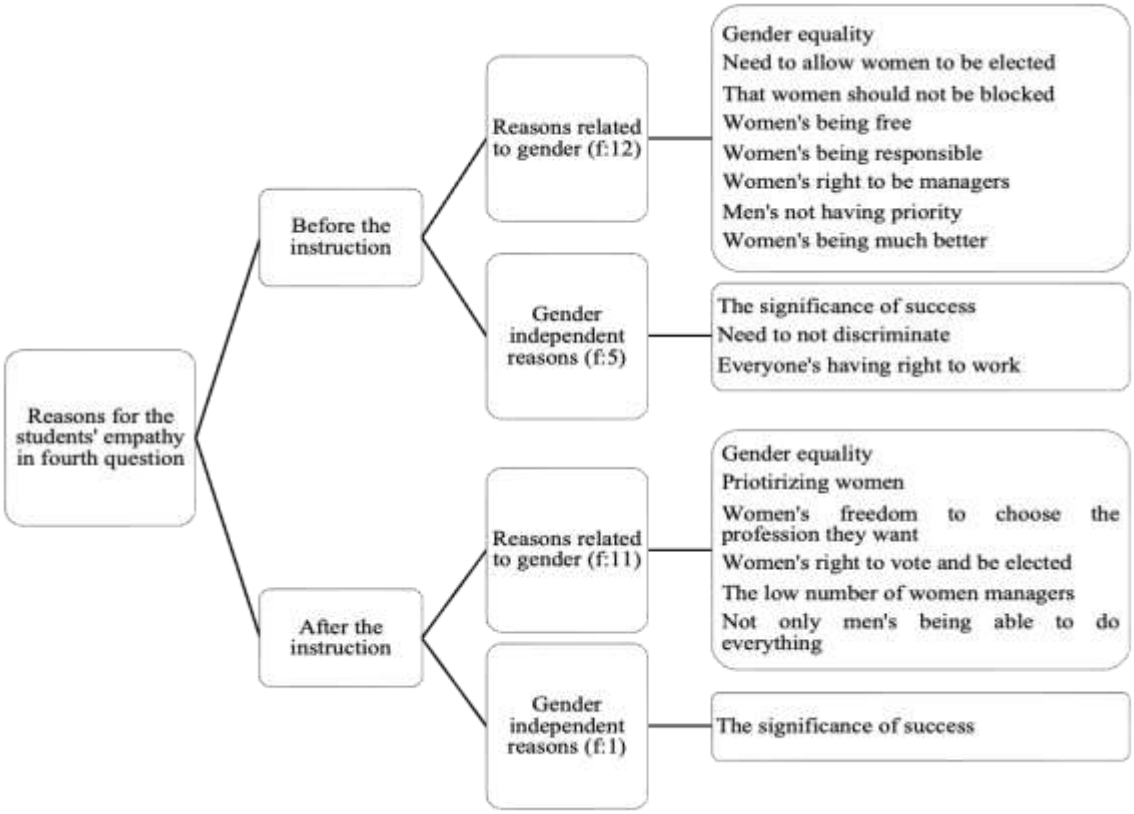


As can be seen in Figure 7, before the instruction, the students mostly associated the reasons they presented for their opinions with gender. Besides, there were students indicating gender independence. Based on this, it could be said that most of the students realized the gender inequality in the case within the fourth question before the instruction, while some did not realize it.

On the other hand, students associated most of their reasons with gender. However, some students presented gender-independent reasons. Compared with the situation before the instruction, it was seen that the reasons given by the students regarding gender did not change in number. Yet, it is thought that the reasons given by students before and after the instruction regarding gender were more numerous. In addition, it was determined that presented gender-independent causes decreased. In line with this, it can be said that students' awareness of gender was at a good level before the instruction, and there was not much change after the instruction. After the instruction, all students stated that they would accept everyone's application regardless of gender if they put themselves in the place of the workplace owner. The reasons they put forward in Figure 8.

Figure 8

Reasons for the Students' Empathy in the Fourth Question



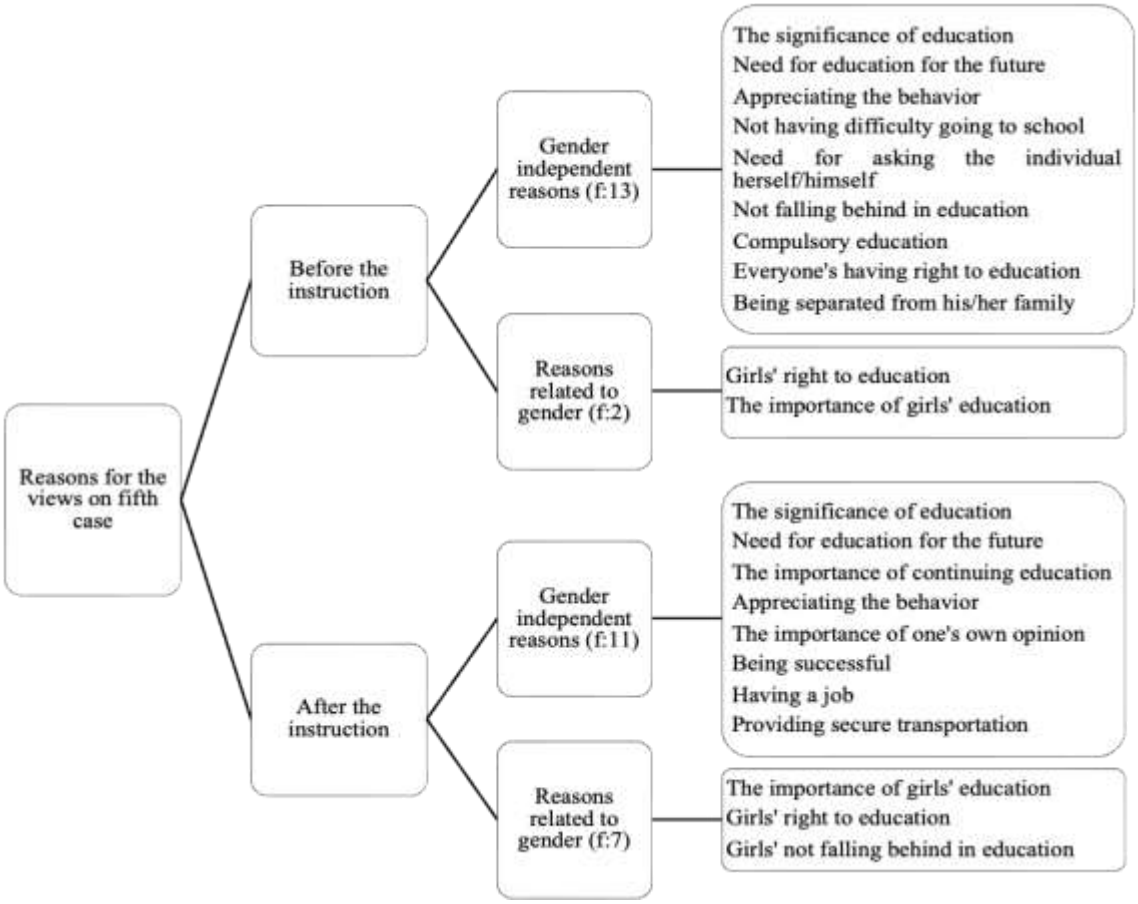
As seen in Figure 8, the reasons given by the students before the instruction were mostly related to the gender issue. However, there were also students who presented their justifications regardless of gender. In the fourth question, it was determined that most of the reasons and justifications presented by the students for the situation in the case were related to the gender issue in general. In addition, it was observed that reasons

and justifications are associated with gender topics as well. Based on this, it could be said that they handled the situation in the case together with the gender topic. However, after the instruction, it was seen that students linked most of their reasons with gender. Moreover, one student presented a gender-independent reason.

Before the instruction, while 19 students stated that they found the sample case in the fifth question correct, one was undecided, and one did not. As for after the instruction, while 20 students found the sample case correct in the fifth question, one student was undecided. They indicated the reasons for these views in Figure 9.

Figure 9

Reasons for the Views on the Fifth Case

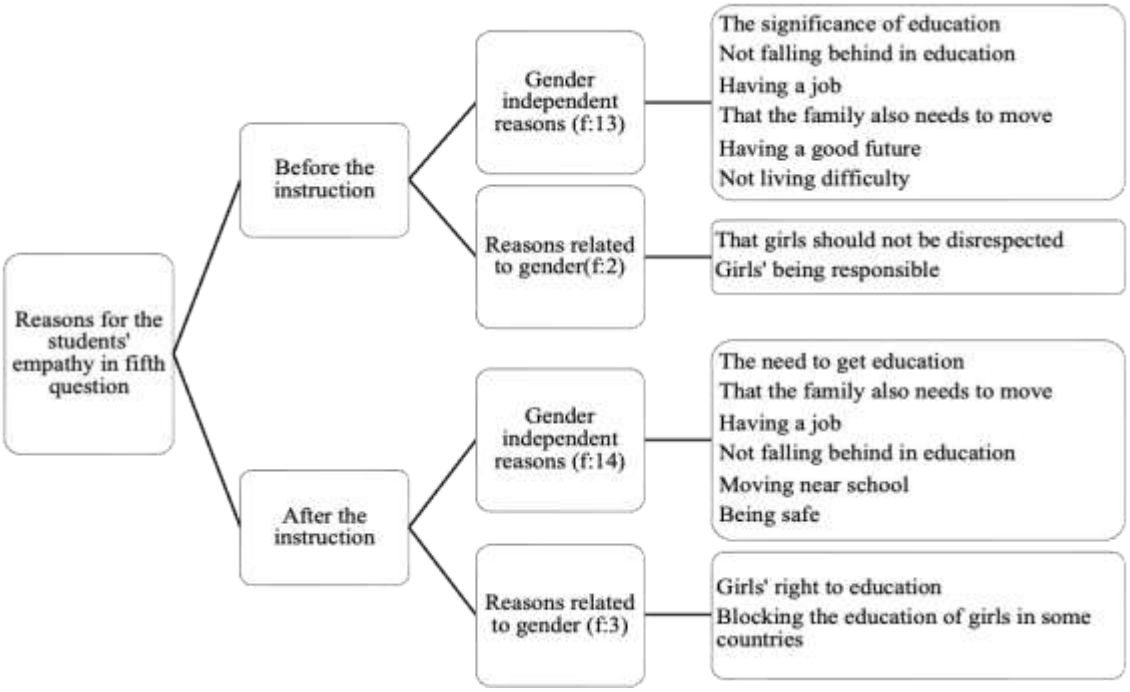


As seen in Figure 9, most of the reasons given by the students before the instruction were about gender-independent issues. However, there were also students who offered reasons about gender. Based on this, it can be said that most of the students did not notice the gender inequality in the case in the fifth question before the instruction, while a small number of students noticed the gender inequality. After the instruction, most of the reasons presented by the students were explained regardless of gender. In addition, it was determined that the reasons associated with the gender issue were also presented. When compared to the situation before the instruction, it was seen that the reasons given by the students for gender increased in number. In line with this, it can be said that students' awareness of gender has improved while questioning the reasons for the situation.

Before the instruction, in the case of the situation, they put themselves in the family's place. 19 students stated that they would behave the same way, one student said the person should be asked first, and one student indicated that the family should go together. On the other hand, after the instruction, when they put themselves in the family's place, 17 students stated that they would act the same way, three students said the family should move together, and one student stated the person should be asked first. The reasons for these views are given in Figure 10.

Figure 10

Reasons for the Students' Empathy in the Fifth Question



As seen in Figure 10, some of the reasons given by the students before the instruction were gender independent, while some of them presented reasons based on gender. In the fifth question, it was determined that most of the reasons and justifications presented by the students for the situation in the case were not related to the gender issue. Based on this, it can be said that most of the students did not deal with the case study on the subject of gender equality. After the instruction, it was determined that most of the students presented reasons regardless of gender. However, it was also observed that there were reasons associated with the gender issue.

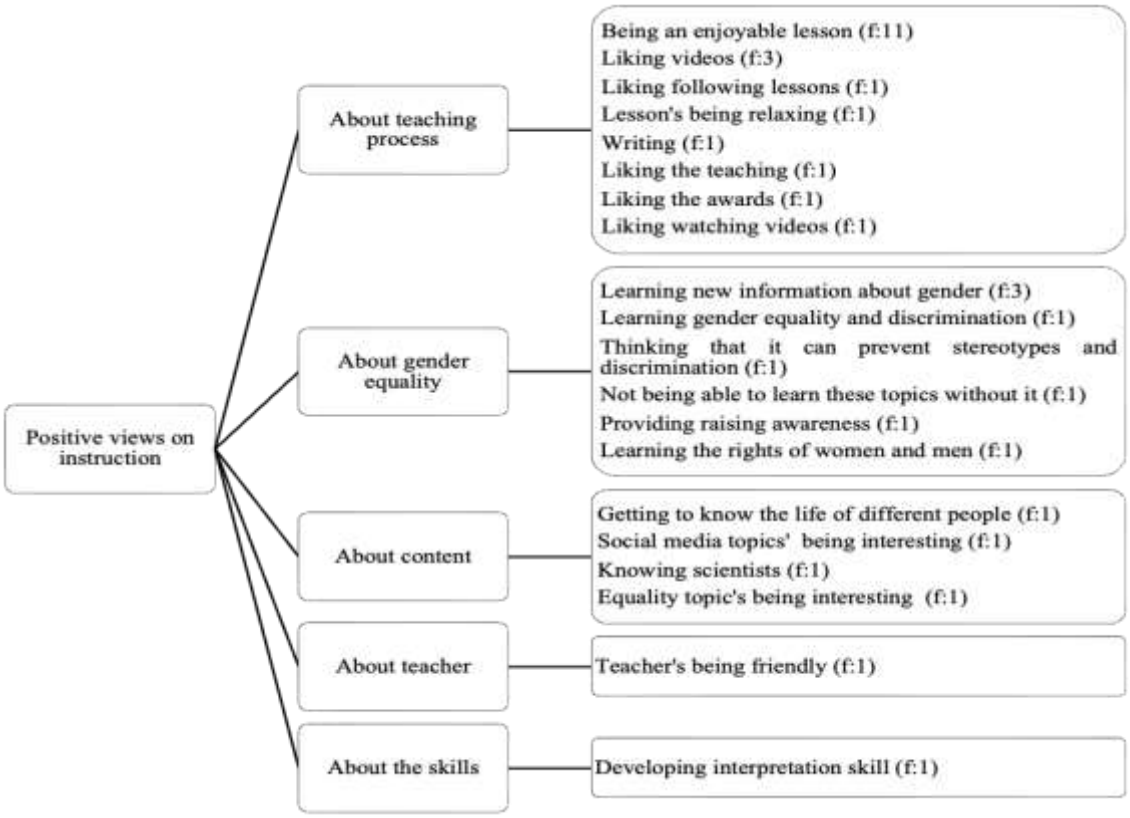
In general, before the instruction, it was seen that the students were aware of gender equality/inequality in some example cases, and in some cases, they were not aware of it. However, it was observed that the reasons and justifications presented for gender equality were less than those presented independently of the issue of gender equality. From this point of view, it could be said that there was a need to improve students' awareness of gender equality before the instruction. In general, when the situation before and after the instruction is compared and the reasons given by the students regarding the case are examined, it can be said that there is an improvement in students' awareness of gender equality after the instruction.

Findings Regarding the Views of Students Towards the Instruction

Students' opinions on instruction were presented under six themes: positive views, negative views on instruction, views on the differences of instruction from other courses, changes in perspective, what was learned at the end of the instructional process, and suggestions for the instructional process. The positive views of the students about the instruction are given in Figure 11.

Figure 11

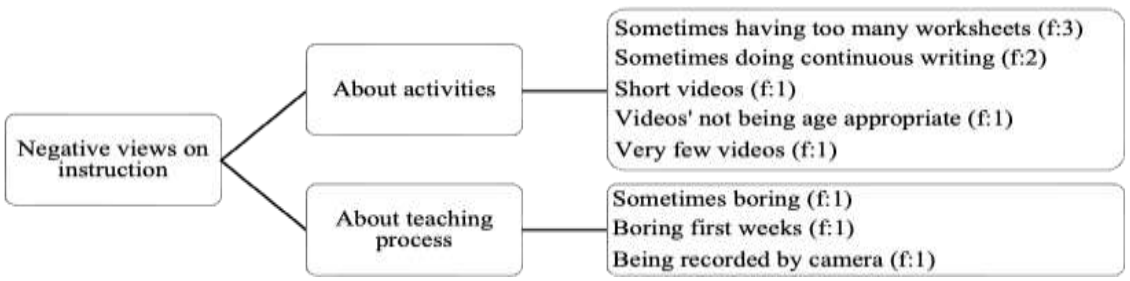
The Positive Views of Students on Instruction



As seen in Figure 11, the students' positive views on instruction were gathered under five themes, which were about the teaching process, gender equality, content, teacher, and skills. Students put forward some views as it being an enjoyable lesson, learning new information about gender topics and learning gender equality and discrimination. The negative views of students on instruction are shown in Figure 12.

Figure 12

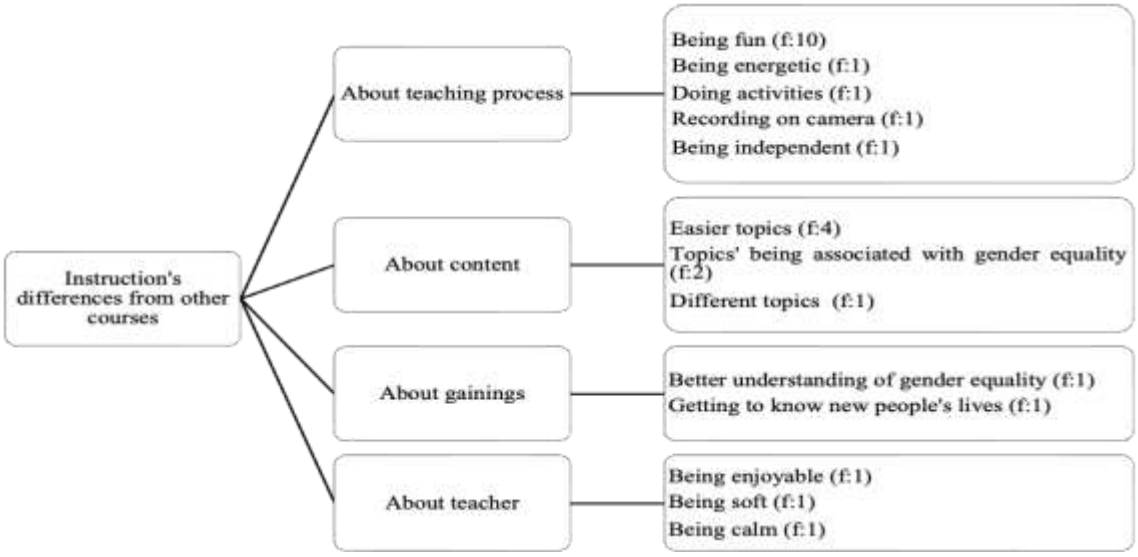
The Negative Views of Students on Instruction



It is observed in Figure 12 that students' negative views on instruction were placed under two themes: activities and the teaching process. Students mostly mentioned sometimes having too many worksheets and sometimes doing continuous writing. Students' views about the instructions' differences from the other courses are presented in Figure 13.

Figure 13

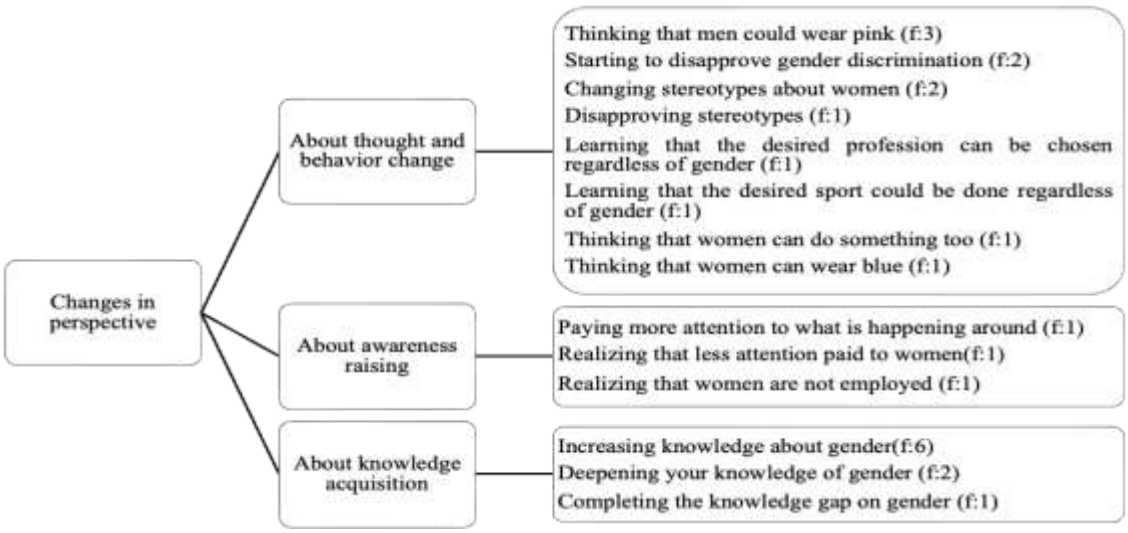
The Views of Students on the Instruction's Differences from Other Courses



Four students stated that the instruction did not differ from other courses. As shown in Figure 13, students' views on the instruction differences from other courses were grouped under four sub-themes: About teaching process, content, gainings, and teacher. The students mostly mentioned that the lesson was fun, the topics were easier, and the topics were associated with gender equality. Changes in students' perspectives are presented in Figure 14.

Figure 14

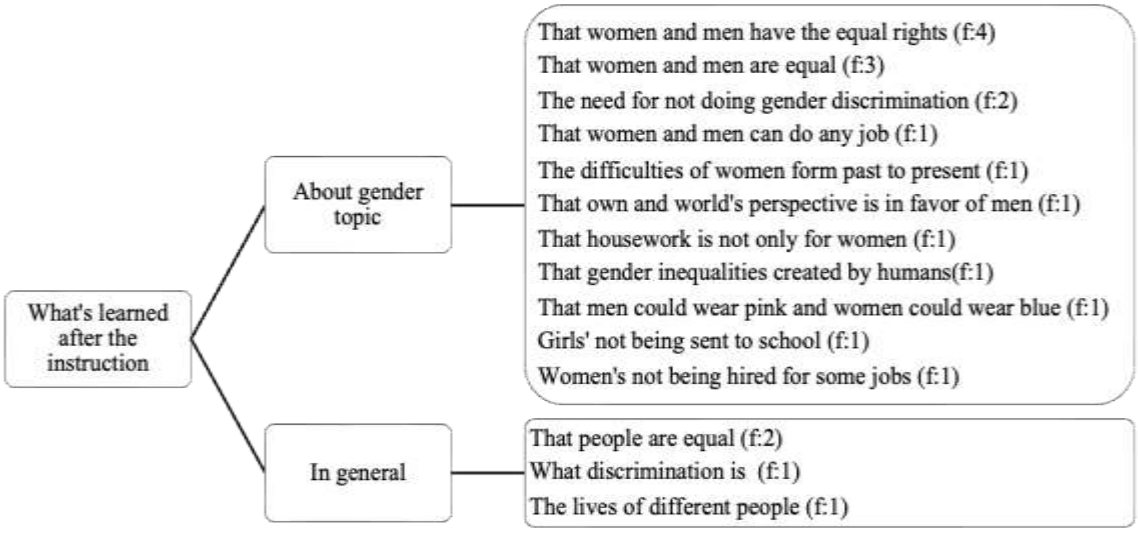
Changes in Students' Perspective



Four students stated that there was no change in their point of view, and they thought similarly before the instruction as well. As can be seen in Figure 14, it was grouped under three sub-themes: about thought and behavior change, awareness raising, and knowledge acquisition. In this context, it was mostly mentioned about the increase in knowledge about gender, thinking that men could wear pink, starting to disapprove of gender discrimination, changing stereotypes about women, and deepening of knowledge about gender. Student views on what was learned at the end of the instructional process are presented in Figure 15.

Figure 15

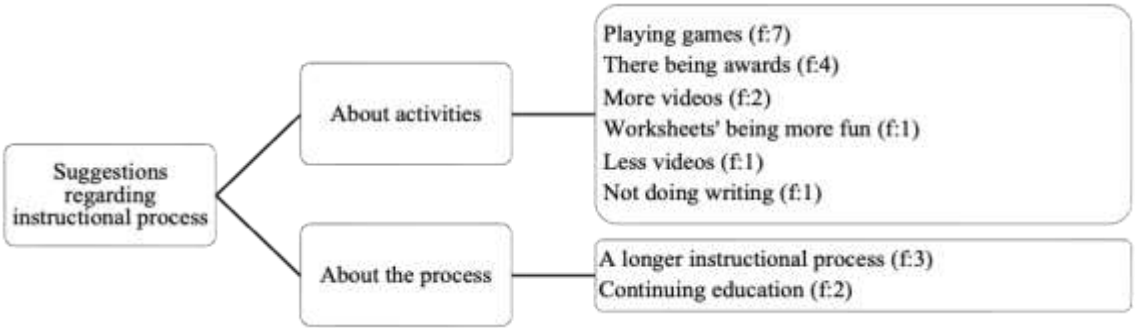
What's Learned After the Instruction



As a result of the instructional process, one student mentioned that they were already familiar with these topics, while two others provided very general responses, stating that they had learned everything. As seen in Figure 15, what was learned at the end of the instruction was gathered under two sub-themes about gender topics and in general. Within this context, they stated that they mostly learned that women and men have equal rights, that women and men are equal, that there should be no gender discrimination, and that people are equal. The suggestions made by the students regarding the instructional process are given in Figure 16.

Figure 16

Suggestions regarding the Instructional Process



As seen in Figure 16, students' suggestions regarding the instructional process were grouped under two sub-themes: about activities and the process. Students mostly indicated that games could be included, awards could be offered, and continuing education should be offered.

Results and Discussion

Before the instruction, when the awareness of the students in the sixth grade of middle school was examined, it was found that the students noticed gender inequality in the cases where there was a clear discourse of gender discrimination when the reasons for their opinions were examined, it was determined that they presented reasons independent of gender. In case studies involving implicit cases of gender equality/inequality, a small number of students did not realize gender equality/inequality; when the reasons for their opinions were examined, it was determined that they mostly presented gender-independent reasons. In line with the answers to the second, third, and fifth questions, it can be said that they did not notice the situation regarding gender equality/inequality; however, in the first and fourth questions, they noticed the gender inequality in the case. In the cases in these questions, it is clear that injustice was done to individuals due to gender stereotypes, but in the cases given in the second and fifth questions, there were situations involving equality, contrary to injustice. As for the case in the third question, there was exclusion instead of injustice. For this reason, they may have associated the students' explanations for the first and fourth questions more with the subject of gender. In addition, when the answers given to the questions were examined, it was determined that while the most reasons for the gender issue were given in the fourth question, in the third question, the least number of reasons for the gender issue were given. Based on this, it can be said that there were deficiencies in gender equality before the instruction and training were needed on this subject. Similar results were obtained with the needs analysis research conducted for this study (Turhan Türkkan et al., 2018a). In similar studies conducted in Turkey, it has been determined that primary and middle school students have stereotypes and negative opinions about gender equality (Turhan Türkkan et al., 2018a; 2018b; Bağçeli Kahraman & Başal, 2011; Kalaycı & Hayırsever, 2014; Kılıç et al., 2014; Yeşil & Balcı Karaboğa, 2021; Yolcu, 2021; Yolcu & Sarı, 2018). As a matter of fact, in different studies on this subject at the middle school level, it is stated that there is a need for education on gender equality (Acar-Erdol & Gözütok, 2017). In line with this, it can be said that it would be beneficial to improve the students' awareness regarding gender equality in the classroom where the instruction was conducted.

In order to determine how interdisciplinary teaching on gender equality in the sixth grade of middle school contributed to students' awareness of gender equality, the reasons given by the students before and after the instruction were examined. In the first, second, third, and fifth questions, it was seen that the reasons given by the students related to gender increased in number compared to the situation before the instruction. Based on this, it could be said that students' awareness of gender has improved while questioning the reasons for the situation. However, in the fourth question, it was seen that the reasons given by the students related to gender did not change in number compared to the situation before the instruction. In the context of the fourth question, it could be inferred that the students' awareness of gender was at a good level before the

instruction while questioning the reasons for the situation, and there was not much change after the instruction. However, in general terms, it can be said that the instruction increased students' awareness of gender equality. In studies conducted at primary and middle school levels, it has been concluded that gender equality education is effective in reducing students' prejudices towards gender, developing concept knowledge about gender, changing their perspectives on gender roles, and raising awareness of gender equality (Akita & Mori, 2022; Brinkman, 2009; Seçgin & Kurnaz, 2015; Yeşil & Balcı Karaboğa, 2021; Yolcu, 2021). In this respect, it could be said that gender equality education positively effects the individual and social context.

When the students' views on the instructional process were analyzed, it was seen that the number of positive opinions about the process was high, and they also presented negative opinions, even a few. However, it has been determined that the instruction differs from other courses, especially in terms of teaching, creating positive changes in the students' perspective, and increasing their learning about the subject. The subject of gender equality must come to the fore in the opinions expressed by the students, and this situation gives rise to the idea that it increases students' awareness of the subject. In this context, it confronts us as an important outcome for students to gain knowledge and awareness about gender equality and to create positive thinking and behavior changes. Based on this, it could be indicated that instruction positively effected students. When the studies on gender equality education and practices are examined, it is seen that education for gender equality creates a positive change in gender attitudes from traditional to egalitarian attitudes (Özcan, 2012), improves gender perception (Altınova & Adıgüzel, 2013), and increases gender sensitivity. (Esen, 2013), improves gender attitudes and raises awareness (Seçgin & Kurnaz, 2015). As seen from the results of these studies, gender equality education positively effects individuals. However, it was determined that some arrangements should be made for the worksheets and videos on the negative opinions of the students and their suggestions to improve the instruction. With respect to this, it can be said that it would be beneficial to improve the teaching activities carried out.

A significant result of this study is also that it is possible to integrate the topic of gender equality with mathematics, science, and ITS courses. From this point of view, it can be said that numerical courses can be a tool for raising awareness about gender equality. Although no study directly integrates gender equality with mathematics courses, when the results of studies that associate mathematics courses with social issues are examined, it has been determined that students have achieved positive outcomes on issues such as equality and social justice (Allen, 2003; Brantlinger, 2007; Gutstein, 2003; Harper, 2017; Johnson, 2005; Koestler, 2010; McNamee, 2013; Turhan Türkkan, 2017; Voss, 2015; Wonnacott, 2011). a study incorporating technology for social justice education concluded that technology can effectively support educational initiatives and actions related to social justice (Mitchell, 2015). Based on this, it is considered that social issues could be included not only in social field courses but also in numerical field courses, and they could be conducted through interdisciplinary instruction.

Considering the positive effects of gender equality education on students, suggestions for practice and future studies were created. These recommendations are presented below:

- As gender equality education increases students' awareness of gender equality, this instructional practice could be extended to the whole country and included in courses' curricula in an interdisciplinary way.
- This study was conducted with middle school sixth-grade students. Similar studies can be carried out at primary education and high school levels.
- This study was carried out for mathematics, science, and ITS courses. In this context, other studies can focus on the integration of different courses.
- This study was carried out with a qualitative method. In other studies, will be conducted in this context, quantitative tools such as scales and tests can be used.
- The case study design was used in this study and was limited to the effects of the instruction on students. In the studies to be carried out in the same context, more detailed studies can be conducted using designs such as action research, ethnography, etc.
- Only students are included as data sources in this study. Stakeholders such as teachers, administrators, and parents can also be included in the studies to be carried out.

Acknowledgements

We would like to thank Çukurova University Scientific Research Projects Coordination Department for its contribution and support to the project with grant number SBA-2018-10028.

Statement of Responsibility

The authors assume full responsibility for the study. In the research, Buket Turhan Türkkan took part in the research design, literature review, implementation, data collection, data analysis and reporting stages. Nihan Arslan Namlı and Betül Karaduman took part in the literature review, implementation, data collection, data analysis and reporting stages of the research. Memet Karakuş took part in the research design, implementation, data analysis and reporting stages of the research. All authors participated in critical review.

Conflicts of Interest

The authors do not have a conflict of interest with any institution or person.

Author Bios:

Buket Turhan Türkkan has a PhD in Curriculum and Instruction at Çukurova University in Adana, Türkiye (2017); MA Curriculum and Instruction at Anadolu University in Eskişehir, Türkiye (2011); BA in Elementary Mathematics Teaching at Eskişehir Osmangazi University in Eskişehir, Türkiye (2008). Turhan Türkkan has studies on curriculum development, mathematics teaching, problem posing, social justice education, gender equality, creativity, interdisciplinary curriculum and teacher education. Currently lecture introduction to education, teaching principles and methods, measurement and evaluation in education, curriculum development in education, classroom management, inclusive education courses. Turhan Türkkan is assistant

professor at Çukurova University, Faculty of Education, Department of Curriculum and Instruction.

Nihan Arslan Namlı completed her undergraduate education in İhsan Doğramacı Bilkent University at Computer Education and Instructional Technologies Department in 2013 and her master's degree without thesis in 2014 in the same department. Arslan Namlı completed her master's degree with thesis in Çukurova University Computer Education and Instructional Technologies Department in 2016 and received her doctorate from the Educational Sciences department at the same university in 2021. Arslan Namlı started her academic life at Çukurova University in 2014 and has been continuing her academic career at Iskenderun Technical University since 2020. Arslan Namlı works on instructional technologies, robotics, computational thinking skills and gamification and has published international book chapters, articles and papers in these fields.

Betül Karaduman has a PhD in Curriculum and Instruction at Çukurova University in Adana (2016), Türkiye; MA Science Teaching at the University Çukurova, Adana, Türkiye (2008); BA in Chemistry Education at the University of Gazi, Ankara, Türkiye (2005). Areas of interest in scientific research are Curriculum development, Science Teaching, and Instructional Technologies. Currently lecture educational science, science curriculum, and out of school learning courses and research assistant at the Department of Mathematics and Science Education in Çukurova University.

Memet Karakuş has has PhD and MA in Curriculum and Instruction at Çukurova University in Adana, Türkiye (2000, 2004); BA in Primary Education at Trakya University, Edirne, Türkiye (1994). Karakuş has studies on curriculum development, social sciences teaching, creativity, curriculum evaluation, interdisciplinary curriculum and primary school teaching. Currently lecture teaching principles and methods, measurement and evaluation in education, curriculum development in education and curriculum and instruction in the interdisciplinary context. Karakuş is associate professor at Çukurova University, Faculty of Education, Department of Curriculum and Instruction.

References

Acar-Erdol, T. (2019). *Toplumsal cinsiyet eşitliği eğitim programının hazırlanması, uygulanması ve değerlendirilmesi* [Yayımlanmamış doktora tezi]. Ankara Üniversitesi.

Acar-Erdol, T., & Gözütok, F. D. (2017). Ortaöğretim öğrencileri için toplumsal cinsiyet eşitliği öğretim programı ihtiyaç analizi: (Bir Anadolu lisesi örneği). *Eğitim ve Bilim, 42*(2017), 36-65.

Acar-Erdol, T., Özen, F., & Toraman, Ç. (2019). Türkiye'deki eğitim fakültesi öğrencilerinin toplumsal cinsiyet eşitliğine yönelik görüşleri. *Kırşehir Eğitim Fakültesi Dergisi, 20*(2), 793-844.

Akita, S., & Mori, K. (2022). Akita, S., & Mori, K. (2022). How implicit image of woman changed in Japanese sixth-grade children after a gender equality education lesson. *The Journal of Social Studies Research, 46*(2), 153-159. https://doi.org/10.1016/j.jssr.2021.05.004

Allen, A. F. (2003). *Integrating themes of social justice and equity into a middle school mathematics classroom* [Unpublished doctoral dissertation]. Indiana University.

Altınova, H. H., & Adıgüzel, Ö. (2013). Yaratıcı drama yöntemiyle verilen toplumsal cinsiyet eğitimi programının değerlendirilmesi. *Yaratıcı Drama Dergisi, 8*(16), 15-32.

Apple, M. W. (2017). *Eğitim toplumu değiştirebilir mi?* Ş. Çınkır (Çev. Ed.). Ankara: Anı Yayıncılık. (Orijinal çalışma 1988 yılında yayımlanmıştır.)

Aslan, G. (2015). Öğretmen adaylarının toplumsal cinsiyet algılarına ilişkin metaforik bir çözümleme. *Eğitim ve Bilim, 40*(181), 363-384. https://doi.org/10.15390/EB.2015.2930

Aydemir, S. (2019). *Toplumsal cinsiyet eşitliğine duyarlı medya okuryazarlığı eğitiminin öğretmen adaylarının toplumsal cinsiyet eşitliğine ilişkin tutumlarına etkisi* [Yayımlanmamış doktora tezi]. Gazi Üniversitesi.

Bağçeli Kahraman, P., & Başal, H. A. (2011). Anne eğitim düzeyine göre çocukların cinsiyet kalıpyargıları ile oyun ve oyuncak tercihleri. *e-Journal of New World Sciences Academy, 6*(1), 1335-1357.

Brantlinger, A. M. (2007). *Geometries of inequality: Teaching and researching critical mathematics in a low-income urban high school* [Unpublished doctoral dissertation]. Northwestern University.

Brinkman, B. G. (2009). *Evaluation of the fair program: Teaching gender equality to children* [Unpublished doctoral dissertation]. Colorado State University.

Çelik, T., Aydoğan Yenmez, A., & Gökçe, S. (2019). Ortaokul matematik ders kitaplarındaki dilsel ve görsel metinlerin toplumsal cinsiyet rollerine göre incelenmesi. *Ondokuz Mayıs Üniversitesi Eğitim Fakültesi Dergisi, 38*(2), 206-224.

Cope, D. G. (2014). Methods and meanings: credibility and trustworthiness of qualitative research. *Oncology Nursing Forum, 41*(1), 89-91.

Drake, S. M., & Reid, J. L. (2020). 21st century competencies in light of the history of integrated curriculum. *Front. Educ. 5*(122). https://doi.org/10.3389/feduc.2020.00122

Esen, Y. (2013). Hizmet öncesi öğretmen eğitiminde toplumsal cinsiyet duyarlılığını geliştirme amaçlı bir çalışma. *Eğitim ve Bilim, 38*(169), 280-295.

Gonzalez, L. (2009). Teaching mathematics for social justice: reflections on a community of practice for urban high school mathematics teachers. *Journal of Urban Mathematics Education, 2*(1), 22-51.

Gözütok, F. D., Toraman, Ç., & Acar-Erdol, T. (2017). Toplumsal cinsiyet eşitliği ölçeğinin (TCEÖ) geliştirilmesi. *İlköğretim Online, 16*(3), 1036-1048.

Güney, N. (2016). 6. sınıf Türkçe ders kitaplarında toplumsal cinsiyet incelemesi. *Turkish Studies, 11*(3), 1229-1248. http://dx.doi.org/10.7827/TurkishStudies.9166

Gutstein, E. (2003). Teaching and learning mathematics for social justice in urban, Latino school. *Journal for Research in Mathematics Education, 34*(1), 37-73.

Gutstein, E. (2007a). Math, maps, and misrepresentations. In W. Au, B. Bigelow & S. Karp (Eds.). *Rethinking our classrooms: Teaching for equity and justice Volume-2* (2nd ed.). (pp. 112-114). Rethinking Schools.

Gutstein, E. (2007b). "And that's just how it starts": Teaching mathematics and developing student agency. *Teachers College Record, 109*(2), 420-448.

Harper, F. K. (2017). Coming to understand the big issues: remaking meaning of social justice through mathematics across the school year. In A. Chronaki (Ed.), Mathematics *Education and Life at Times of Crisis, 2, Proceedings of the Ninth International Mathematics Education and Society Conference* (pp. 513-521). Volos: University of Thessaly Press.

Harris, K. L., Melaas, K., & Rodacker, E. (1999). The Impact of Women's Studies Courses on College Students of the 1990s. *Sex Roles, 40*(11-12), 969-977.

Hellsten, S. K. (2007). From information society to global village of wisdom? The role of ICT in realizing social justice in the developing world. In E. Rooksby & J. Weckert (Eds.), *Information Technology and Social Justice* (pp. 1-28). IGI Global.

İnci, S., & Kaya, V. H. (2022). Eğitimde multidisipliner, disiplinlerarası ve transdisipliner kavramları. *Milli Eğitim, 51*(235), 2757-2772.

Işık Demirhan, E. (2021). Toplumsal cinsiyet eşitliği açısından öğretim programı ve ders kitabı inceleme (4. sınıf insan hakları, yurttaşlık ve demokrasi). *Journal of Innovative Research in Social Studies, 4*(1), 69-82. DOI: https://doi.org/10.47503/jirss.931329

Jacobs, H. H. (1989). Design options for an integrated curriculum. In *Interdisciplinary curriculum: Design and implementation*, H. H. Jacobs (Ed.), pp.13-24. Association for Supervision and Curriculum Development.

Johnson, J. D. (2005). *Instructional implications of social justice pedagogy on the teaching of mathematics* [Unpublished doctoral dissertation]. The Florida State University.

Kalaycı, N., & Hayırsever, F. (2014). Toplumsal cinsiyet eşitliği bağlamında vatandaşlık ve demokrasi eğitimi ders kitabına yönelik bir inceleme ve bu konuya ilişkin öğrenci algılarının belirlenmesi. *Kuram ve Uygulamada Eğitim Bilimleri, 14*(3), 1-26.

Karakuş, E., Mutlu, E., & Diker Coşkun, Y. (2018). Toplumsal cinsiyet eşitliği açısından öğretim programlarının incelenmesi. *İstanbul Üniversitesi Kadın Araştırmaları Dergisi, 2*, 31-54.

Karakuş, M., Turhan Türkkan, B., & Karakuş, F. (2017). Fen bilgisi ve ilköğretim matematik öğretmenlerinin disiplinlerarası yaklaşıma yönelik görüşlerinin belirlenmesi. *İlköğretim Online, 16*(2), 509-524.

Kellogg, P. (2002). Ten chairs of inequality. In B. Bigelow & B. Peterson, (Eds.), *Rethinking Globalization – Teaching for Justice in an Unjust World* (pp.115-117). Rethinking Schools.

Kılıç, A. Z., Beyazova, A., Akbaş, H. M., Zara, A., & Serhatlı, İ. (2014). Okul çağı çocuklarının toplumsal cinsiyet algıları: gündelik yaşam örnekleriyle cinsiyetçiliğin benimsenme durumuna ve esneyebilme olasılığına dair bir araştırma. *Sosyoloji Araştırmaları Dergisi, 17*(2), 122-151.

Koestler, C. (2010). *(Re)envisioning mathematics education: examining equity and social justice in an elementary mathematics methods course* [Unpublished doctoral dissertation]. University of Wisconsin – Madison.

Kollmayer, M., Schultes, M. T., Lüftenegger, M., Finsterwald, M., Spiel, C., & Schober, B. (2020). REFLECT - A teacher training program to promote gender equality in schools. *Frontiers in Education, 5*, 1-8. https://doi.org/10.3389/feduc.2020.00136

Koyuncu Şahin, M., Esen Çoban, A., & Korkmaz, A. (2018). Toplumsal cinsiyet eşitliği ve Türk eğitim sistemindeki yeri: Okul öncesi öğretmen adaylarının gözünden. *Uluslararası Bilimsel Araştırmalar Dergisi, 3*(2), 735-752. http://dx.doi.org/10.21733/ibad.457232

Kükrer, M., & Kıbrıs, İ. (2017). CEDAW öncesi ve sonrası ortaokul Türkçe ders kitaplarında yer alan toplumsal cinsiyet eşitliği faktörünün değerlendirilmesi. *Abant İzzet Baysal Üniversitesi Eğitim Fakültesi Dergisi, 17*(3), 1369-1383.

Laçin Şimşek, C. (2009). How much and how science and technology curriculums and textbooks benefits from history of science? *Elementary Education Online, 8*(1), 129-145.

Leonard, J., Brooks, W., & Barnes-Johnson, J. (2010). The nuances and complexities of teaching mathematics for cultural relevance and social justice. *Journal of Teacher Education, 61*(3), 261-270.

Levine, D., Lowe, R., Peterson, B., & Tenorio, R. (1995). Multiculturalism and antibias education. In D. Levine, R. Lowe, B. Peterson & R. Tenorio (Eds.), *Rethinking schools – An agenda for change* (pp. 5-6). The New Press.

Maya, İ. (2013). Türk eğitim sistemindeki cinsiyet eşitsizliklerinin AB ülkeleri ile karşılaştırılması. *Eğitim ve Bilim, 38*(168), 69-84.

Mayberry, M. (1998). Reproductive and resistant pedagogies: The comparative roles of collaborative learning and feminist pedagogy in science education. *Journal of Research in Science Teaching: The Official Journal of the National Association for Research in Science Teaching, 35*(4), 443-459.

McBee, R. H. (1996). Can controversial topics by taught in the early grades. *Social Education, 60*(1), 38-41.

McNamee, R. (2013). *Teaching social justice mathematics in a privileged setting* [Unpublished doctoral dissertation]. Montclair State University.

Millar, V. (2020). Trends, issues and possibilities for an interdisciplinary STEM curriculum. *Science & Education, 29*(4), 929-948.

Miner, B. (1995). Taking multicultural, anti-racist education seriously: an interview with Enid Lee. In D. Levine, R. Lowe, B. Peterson & R. Tenorio (Eds.), *Rethinking schools – An agenda for change* (pp. 9-16). The New Press.

Ministry of National Education. (2018). *Information technologies and software course curriculum (Secondary school 5th and 6th grades).* Ankara.

Mitchell, L. (2015). *The development of praxis, empowerment and capacity for social justice action among youth participants in a technology- embedded, online social justice education program* [Unpublished master's thesis]. Queen's University.

Munkebye, E., Scheie, E., Gabrielsen, A., Jordet, A., Misund, S., Nergård, T., & Øyehaug, A. B. (2020). Interdisciplinary primary school curriculum units for sustainable development. *Environmental Education Research, 26*(6), 795-811.

Noyes, A. (2007). *Rethinking school mathematics*. Paul Chapman Publishing.

Öngen, B., & Aytaç, S. (2013). Üniversite öğrencilerinin toplumsal cinsiyet rollerine ilişkin tutumları ve yaşam değerleri ilişkisi. *Istanbul Journal of Sociological Studies, 48*(2), 1-18.

Özcan, A. (2012). *Toplumsal cinsiyet eğitiminin üniversite öğrencilerinin toplumsal cinsiyet rol tutumları üzerine etkisi* [Yayımlanmamış doktora tezi]. Erciyes Üniversitesi.

Patton, M. Q. (2002). *Qualitative research & evaluation methods* (3rd ed.). Sage Publications.

Pedretti, E. (1999). Decision making and sts education: exploring scientific knowledge and social responsibility in schools and science centers through an issue-based approach. *Journal of School Science and Mathematics, 99*(4), 174-181.

Peterson, B. (2002). Planting seeds of solidarity. In B. Bigelow & B. Peterson, (Eds.), *Rethinking globalization – Teaching for justice in an unjust world* (pp. 18-28). Rethinking Schools Publication.

Polat, S. (2010). *Eğitimde ve sınıf içi süreçlerde cinsiyetçi yaklaşımlar* [Yayımlanmamış yüksek lisans tezi]. İstanbul Üniversitesi.

Punch, K. F. (2005). *Sosyal araştırmalara giriş – nitel ve nicel yaklaşımlar.* (D. Bayrak, H. B. Arslan & Z. Akyüz, Çev.). Siyasal Kitabevi. (Orijinal çalışma 2005 yılında yayımlanmıştır)

Roth, W. M. (2007). Toward a dialectical notion and praxis of scientific literacy. *Journal of Curriculum Studies, 39*(4), 377-398.

Schniedewind, N., & Davidson, E. (2006). *Open minds to equality: A sourcebook of learning activities to affirm diversity and promote equity.* Rethinking Schools.

Seçgin, F., & Kurnaz, Ş. (2015). Sosyal bilgiler dersinde toplumsal cinsiyet etkinliklerinin öğrencilerin algı ve tutumlarına etkisi. *Uluslararası Türk Eğitim Bilimleri Dergisi, 3*(5), 24-38.

Şener Özel, F. (2019). *Eğitim fakültelerinin lisans programlarına yönelik toplumsal cinsiyet eğitimi dersi öğretim programının tasarlanması, uygulanması ve değerlendirilmesi* [Yayımlanmamış yüksek lisans tezi]. Düzce Üniversitesi.

Sis-Çelik, A., Pasinlioğlu, T., Tan, G., & Koyuncu, H. (2013). Üniversite öğrencilerinin cinsyet eşitliği tutumlarının belirlenmesi. *F. N. Hemşirelik Dergisi, 21*(3), 181-186.

Stake, J. E. (2006). Pedagogy and student change in the women's and gender studies classroom. *Gender and Education, 18*(2), 199-212. http://dx.doi.org/10.1080/09540250500380687

Stake, J. E., & Hoffmann, F. L. (2001). Changes in student social attitudes, activism, and personal confidence in higher education: The role of women's studies. *American Educational Research Journal, 38*(2), 411-436. http://dx.doi.org/10.3102/00028312038002411

Tarman, B., Baytak, A., & Duman, H. (2015). Teachers' views on an ICT reform in education for social justice. *Eurasia Journal of Mathematics, Science and Technology Education, 11*(4), 865-874.

Thomsen, C. J., Basu, A. M., & Reinitz, M. T. (1995). Effects of women's studies courses on gender-related attitudes of women and men. *Psychology of Women Quarterly, 19*(3), 419-426. http://dx.doi.org/10.1111/j.1471-6402.1995.tb00084.x

Tokuş, K. (2018). *Ortaokul fen bilimleri ders kitaplarının bilim tarihi kullanımı açısından incelenmesi* [Yayımlanmamış yüksek lisans tezi]. Trakya Üniversitesi.

Topçu, M. S., Muğaloğlu, E. Z., & Güven, D. (2014). Fen eğitiminde sosyobilimsel konular: Türkiye örneği. *Kuram ve Uygulamada Eğitim Bilimleri, 14*(6), 1-22.

Turhan Türkkan, B. (2017). *Sosyomatematiksel konularla bütünleştirilmiş matematik öğretimi: Sosyal adalet ve eşitlik değerlerine ilişkin farkındalık ile problem kurma becerisi geliştirmeye yönelik bir eylem araştırması* [Yayımlanmamış doktora tezi]. Çukurova Üniversitesi.

Turhan Türkkan, B., Karaduman, B., Arslan Namlı, N., & Karakuş, M. (2018a, 11-13 Ekim). *Ortaokul öğrencilerinin toplumsal cinsiyet eşitliğine yönelik farkındalıklarının belirlenmesi: Bir ihtiyaç analizi çalışması* [Paper presentation]. 6. Uluslararası Eğitim Programları ve Öğretim Kongresi, Kars, Türkiye.

Turhan Türkkan, B., Karakuş, M., Arslan Namlı, N., & Karaduman, B. (2018b, September 4-6). *Toplumsal cinsiyet eşitliği eğitimine yönelik bir ihtiyaç analizi çalışması: Öğretmenlerden yansımalar* [Paper presentation]. 6th International Multidisciplinary Congress of Eurasia, Rome, Italy.

TURKSTAT. (2021a, May 17). *İstatistiklerle gençlik, 2020*. TÜİK. https://data.tuik.gov.tr/Bulten/Index?p=Istatistiklerle-Genclik-2020-37242

TURKSTAT. (2021b, October 26). *Araştırma-geliştirme faaliyetleri araştırması, 2020*. TÜİK. https://data.tuik.gov.tr/Bulten/Index?p=Arastirma-Gelistirme-Faaliyetleri-Arastirmasi-2020-37439

TURKSTAT. (2021c, March 5). *İstatistiklerle kadın, 2020*. TÜİK. https://data.tuik.gov.tr/Bulten/Index?p=Istatistiklerle-Kadin-2020-37221

Vefikuluçay, D., Zeyneloğlu, S., Eroğlu, K., & Taşkın, L. (2007). Kafkas Üniversitesi son sınıf öğrencilerinin toplumsal cinsiyet rollerine ilişkin bakış açıları. *Hemşirelik Yüksekokulu Dergisi, 2007*, 26–38.

Vefikuluçay Yılmaz, D., Zeyneloğlu, S., Kocaöz, S., Kısa, S., Taşkın, L., & Eroğlu, K. (2009). Üniversite öğrencilerinin toplumsal cinsiyet rollerine ilişkin görüşleri. *Uluslararası İnsan Bilimleri Dergisi, 6*(1), 775-792.

Voss, R. N. (2015) *Teaching mathematics for social justice within a Victorian public school for year nine mixed ability classes* [Unpublished master's thesis]. Curtin University.

Wager, A. A. (2008). *Developing equitable mathematics pedagogy* [Unpublished doctoral dissertation]. University of Wisconsin – Madison.

Wonnacott, V. (2011). *Teaching mathematics for social justice and its effects on affluent students* [Unpublished master's thesis]. University of Toronto.

World Economic Forum (2019, December 16). *Global gender gap report 2020.* WEF. https://www3.weforum.org/docs/WEF_GGGR_2020.pdf

World Economic Forum (2021, March 30). *Global gender gap report 2021*. WEF. https://www3.weforum.org/docs/WEF_GGGR_2021.pdf

Yeşil, A., & Balcı Karaboğa, F. A. (2021). Ortaokul öğrencilerinde toplumsal cinsiyet eşitliği farkındalığı. *Eğitimde Nitel Araştırmalar Dergisi, 28*, 51-78. https://doi.org/10.14689/enad.28.3

Yıldırım, A., & Şimşek, H. (2008). *Sosyal bilimlerde nitel araştırma yöntemleri* (7. bs.). Seçkin Yayıncılık.

Yolcu, E. (2021). *Toplumsal cinsiyet eşitliği duyarlılığının geliştirilmesi: İlkokul dördüncü sınıflarla yapılan bir eylem araştırması* [Yayımlanmamış doktora tezi] Çukurova Üniversitesi.

Yolcu, E., & Sarı, M. (2018, September 4-6). *Gender attitudes: A study to determine the case for 4th graders* [Paper presentation]. VI. International Multidisciplinary Congress of Eurasia, Rome, Italy.

Appendix: Views Form on Gender Equality

1) Ayşe and Ali are seventh grader twins. Both wants to be computer engineers after graduated from high school. However, her family doesn't want Ayşe to be an engineer. They do not find it right for women to be engineer so they hinder Ayşe. On the contrary, his family supports Ali's wish for being an engineer as he is a male. In this case, regarding family's thoughts and behaviors;

a) I find it right.

b) I do not find it right

c) I am undecided

Because;……

……

What would you do if you were in the family's place? Why?

……

……

2) Ozan is a 15-year-old boy. Ozan's mother and father are teachers. Ozan's parents are usually very tired when they come home. For this reason, Ozan helps his family in preparing the meal, setting the table and washing the dishes in order to prevent his parents from getting tired more. He helps his family as much as he can in house cleaning on weekends. In this case, about Ozan's helping his family in households;

a) I find it right.

b) I do not find it right.

c) I am undecided.

Because;……

……

What would you do if you were in Ozan's place? Why?

……

……

3) Damla is a middle school sixth grade girl who loves football very much. As she likes to play football, she plays during breaks and in her spare time. However, her girlfriends in the class criticize Damla's playing football and exclude her. In this case, regarding the behavior of Damla's girlfriends towards Damla;

a) I find it right.

b) I do not find it right.

c) I am undecided.

Because;……

……

What would you do if you were Damla's girlfriends? Why?

……

……

4) Fatma is a very successful employee at her job. Fatma is 32 years old and has two children. Fatma successfully manages her home life and business life, and does her stuff at work without slacking off. A manager will be determined at Fatma's workplace this

month. Along with Fatma, two female employees and four male employees applied to become managers. However, the owner of the workplace where they work did not accept the applications of female employees, stating that women could not be managers, and only accepted applications from male employees. In this case, regarding the thoughts and behaviors of the owner of the workplace;

a) I find it right.

b) I do not find it right.

c) I am undecided.

Because;……
……

What would you do if you were the owner of the workplace? Why?
……
……

5) Aslı is a 16-year-old girl who continues her high school education. The village where Aslı lives is far from the nearest town where there is a high school, and in winter the roads are very dangerous due to snow. Aslı's family supports the education of girls. However, they think that Aslı is having difficulties because her high school is far away. Her family, thinking that Aslı would not fall behind in education, bought a house in the district where the high school is located and decided to live with her grandmother there. In this case, this decision of Aslı's family;

a) I find it right.

b) I do not find it right.

c) I am undecided.

Because;……
……

What would you do if you were in Aslı's family's place? Why?

……
……

Journal of Theoretical Educational Science, 17(1), 202-235, January 2024
Kuramsal Eğitimbilim Dergisi, 17(1), 202-235, Ocak 2024
[Online]: http://dergipark.org.tr/akukeg
DOI number: http://doi.org/10.30831/akukeg.1358837

The Effect of Digital Game-Based and Different Education Programs on Phonological Awareness Skills of 60-72 Months-Old Children*

Dijital Oyun Tabanlı ve Farklı Eğitim Programlarının 60-72 Aylık Çocukların Ses Bilgisel Farkındalık Becerilerine Etkisi

Mehmet Oğuz GÖLE** **Zeynep Fulya TEMEL*****

Received: 12 September 2023 **Research Article** **Accepted:** 30 October 2023

ABSTRACT: This study examines the effects of digital game-based and different educational programs on the phonological awareness skills of 60-72-month-old children. The study group of the research consists of 60-72-month-old children studying in kindergartens affiliated with primary schools. The study group was divided into three experimental groups and one control group. There were 22 children in experimental group 1, 16 in experimental group 2, 17 in experimental group 3, and 17 in the control group, totaling 72 children. In the study, a phonological awareness education program was applied to experimental group 1, a digital game-supported phonological awareness education program was applied to experimental group 2, a digital game-based phonological awareness education program was applied to experimental group 3, and no intervention was applied to the control group. A quasi-experimental design was used as the research model. The Study used the Early Literacy Skills Assessment Tool (ELSA) and the Early Literacy Test for Preschoolers (ELTKC) as data collection tools. A pre-test was administered to the study and control groups, and after the pre-test, educational programs were applied to the experimental groups for ten weeks. After the educational programs were implemented, a post-test was administered to the experimental and control groups, and a retention test was administered to the experimental groups. According to the results of the study, it was determined that there was a significant difference between the experimental groups and between the experimental groups and the control group.

Keywords: Phonological awareness, digital game, educational program, preschool education.

ÖZ: Bu araştırmanın amacı dijital oyun tabanlı ve farklı eğitim programlarının 60-72 aylık çocukların ses bilgisel farkındalık becerilerine etkisinin incelenmesidir. Araştırmanın çalışma grubunu ilkokula bağlı anasınıflarında eğitim alan 60-72 aylık çocuklar oluşturmaktadır. Araştırmada çalışma grubu üç deney ve bir kontrol grubu olmak üzere dört farklı gruba ayrılmıştır. Deney grubu 1'de 22, deney grubu 2'de 16, deney grubu 3'de 17 ve kontrol grubunda 17 çocuk olmak üzere toplam 72 çocuk bulunmaktadır. Araştırmada deney 1 grubuna; ses bilgisel farkındalık eğitim programı uygulanmış, deney 2 grubuna dijital oyun destekli ses bilgisel farkındalık eğitim programı uygulanmış, deney 3 grubuna dijital oyun tabanlı ses bilgisel farkındalık eğitim programı uygulanmış, kontrol grubuna ise hiçbir müdahalede bulunulmamıştır. Araştırmanın modeli olarak yarı deneysel desen kullanılmıştır. Araştırmada veri toplama araçları olarak Erken Okuryazarlık Becerilerini Değerlendirme Aracı (EOBDA) ve Anasınıflarına Yönelik Erken Okuryazarlık Testi (EROT) kullanılmıştır. Çalışma gruplarına ve kontrol gruplarına ön test uygulanmış ve ön test sonrasında deney gruplarına on hafta süreyle eğitim programları uygulanmıştır. Eğitim programları uygulandıktan sonra deney gruplarına ve kontrol grubuna son test, deney gruplarına kalıcılık testi yapılmıştır. Araştırmanın sonuçlarına göre deney grupları arasında ve deney grupları ile kontrol grubu arasında anlamlı farklılık olduğu tespit edilmiştir.

Anahtar kelimeler: Ses bilgisel farkındalık, dijital oyun, eğitim program, okul öncesi eğitim.

* It was produced from the doctoral dissertation titled "Digital Game-Based, Supported Training Programs for Phonological Awareness Skills and the Effect of Phonological Awareness Training Program on the Phonological Awareness Skills of 60-72 Month-Old Children".

** *Corresponding Author*: Asst. Prof. Dr., Afyon Kocatepe University, Afyonkarahisar, Türkiye, mogole@aku.edu.tr, https://orcid.org/0000-0003-2826-1790
*** Prof. Dr., Gazi University, Ankara, Türkiye, ftemel@gazi.edu.tr, https://orcid.org/0000-0002-5375-3503

Citation Information
Göle, M. O., & Temel, Z. F. (2024). The effect of digital game-based and different education programs on phonological awareness skills of 60-72 months-old children. *Kuramsal Eğitimbilim Dergisi [Journal of Theoretical Educational Science]*, *17*(1), 202-235.

ISSN: 1308-1659

Quality preschool education affects children's school success in primary school and further education levels. The most important factors affecting literacy success in primary school are quality early literacy education and phonological awareness education given in the preschool period (Bentin et al., 1991; Castles & Coltheart, 2004; Goswami, 2003; Justice & Vukelich, 2008; Kostelnik et al. al., 2019; Kozminsky & Kozminsky, 1995). Phonological awareness is defined as an individual's implicit and explicit sensitivity to the sounds in the structure of spoken language (Pullen & Justice, 2003). Phonological awareness is a starting point for understanding the relationship between a sound and the letters in the alphabet that represent it. It starts at about age three and develops gradually over the years. This skill begins with word and syllable awareness and continues with phoneme awareness, progressing toward the smallest phoneme in the word. It is also stated that phonological awareness occurs when children can hear and notice the boundaries of words (Christie et al., 2011; Morrison & Wilcox, 2012). Phonological awareness skills: realizing that the sentence consists of words, noticing rhyming words and producing rhyming words, separating the word into syllables, determining and matching the first/last sound of the word, determining the word formed by adding sound and syllable to the word, determining the word formed when the sound and syllable are removed from the word, the sounds forming the word and syllable speaking skills form phonological awareness skills (Hempenstall, 2003; National Reading Panel, 2000; Phillips et al., 2008; Yopp & Yopp, 2000).

Phonological awareness studies contribute to separating and combining sounds, spelling, decoding, and reading during the reading process. Writing affects skills such as understanding the principle of the alphabet, coding, understanding the relationship between letters and sounds, and writing in accordance with the letters represented by sounds (Busink, 1997; Cardoso-Martins et al., 2011; Castles & Coltheart, 2004; De Jong, 2007; Martins & Silva, 2001; Schuele & Boudreau, 2008; Torgesen & Mathes, 1998). High-quality phonological awareness education, which includes quality educational environments, methods, and techniques and the implementation process that children are exposed to in the preschool period, enables children to be successful in reading and writing skills in the future, and the effect of this success continues in the first and second grades of primary school (Casalis & Colé, 2009; Hogan et al., 2005). In the studies, it has been determined that enriching the educational environment and learning centers with visual stimuli related to phonological awareness, using visual and audio-sensory cards, and authentic materials supports the development of phonological awareness skills (Chambers et al., 2016; Franc & Subotic, 2015; Sucena et al., 2023). The studies determined that children's rhyme, syllable, and sound awareness skills were supported by the use of children's books that comply with content standards and interactive book reading methods (Elmonayer, 2013; Lefebvre et al., 2011; Lenhart et al., 2022; Mihai et al., 2015; Stadler & McEvoy, 2003; Symmonds, 2020). As a different method in phonological awareness education, poetry-based education programs in which poems are used for children have also been found to improve phonological awareness positively (Lennox, 2014; Lim & Chew, 2018; Nichols et al., 2018). It has been determined that word games played by changing the sounds and syllables in the word, visual cards, and games played using different materials positively affect phonological awareness (Arias, 2023; Luna, 2021). In phonological awareness studies using rhymes and finger games, rhyme awareness provides the development of

skills to combine sounds and syllables (Bolduc & Lefebvre, 2012; Harper, 2011; Incognito & Pinto, 2023; Redig, 2018). Singing children's songs, writing songs with children, and practicing rhythm improves syllable awareness, rhyme, sound awareness, and sound combining skills (Degé & Schwarzer, 2011; Del Egido, 2023; Moritz et al., 2013; Patria, 2023; Rowe et al., 2023).

Many different materials are used in the learning environment for children's phonological awareness skills. Nowadays, tablets and smartphones are essential to children's daily lives, and digital games are among children's game preferences and attract attention (Mertala & Meriläinen, 2019; Rideout, 2011; Shuler, 2012). Therefore, it is seen that digital games are essential in the education process in terms of increasing the diversity of materials and a qualified phonological awareness educational environment for children (Bennett & Parise, 2014; Saracho & Spodek, 2002; Selwyn, 2016; Sheridan et al., 2011; Zevenbergen, 2007). Digital game-based learning has emerged by incorporating digital games into the educational process. Digital game-based learning refers to learning that takes place through digital games. It encompasses learning through playing games or developing games (Becker, 2017). Using digital games in the educational process; critical thinking, problem-solving, knowledge construction (Moyer et al., 2002), using digital games in the educational process; critical thinking, problem-solving, knowledge construction (Moyer et al., 2002), matching, classifying, sorting, counting and coding (Pila et al., 2019; Stephen & Plowman, 2014), cognitive skills such as discovery, creativity, as well as acting autonomously, cooperation, intrinsic motivation (Arnott, 2013; Ferrari & Addessi, 2014; Lennon et al., 2022; McClarty et al., 2012; Ryan et al., 2006; Xiong et al., 2022), it also supports phonological awareness skills, one of the early literacy skills, as well as social skills such as. Studies conducted using digital games have shown contributions to the development of subdimensions of phonological awareness, such as syllable awareness, segmentation of words into syllables (Da Silva et al., 2022; Elimelech & Aram, 2020), syllable merging, syllable deletion from words (Amorim et al., 2020), rhyme detection, and rhyme matching (Goffredo et al., 2016; Puolakanaho et al., 2003; Van Goch et al., 2017), initial sound identification, matching words with the same initial sounds, sound deletion from words, blending sounds and syllables (Arnold et al., 2021; Cameron, 2023; Li et al., 2020; Sá et al., 2022), word segmentation into sounds, and letter knowledge (Kartal & Terziyan, 2016), as well as coding skills (Weiss et al., 2022).

Research has identified various methods, techniques, applications, and educational programs implemented to enhance children's phonological awareness skills. These studies have found that educational programs utilizing poetry, nursery rhymes, children's songs, and books, as well as gamified activities, support children's phonological awareness skills (Arias, 2023; Bayraktar & Temel, 2014; Del Egido, 2023; Incognito & Pinto, 2023; Rowe et al., 2023). Furthermore, research conducted on phonological awareness has concluded that digital games have the potential to enhance children's syllable, rhyme, and phonological awareness skills (Arnold et al., 2021; Kartal & Terziyan, 2016; Oliva-Maza et al., 2021; Vinter et al., 2022). When we look at the research above, the effectiveness of digital game-based, digital game-supported, and phonological awareness studies implemented with different methods was compared with the control group. However, few studies compare the effectiveness of various educational programs among themselves and the control group. The research conducted

is essential in terms of being comprehensive and trying to determine the effectiveness of various educational programs for phonological awareness (digital game-based, digital game-supported, educational programs where different methods are applied) by comparing them between the programs and the control group. It is also thought to be a guide in determining and selecting phonological awareness educational programs for children.

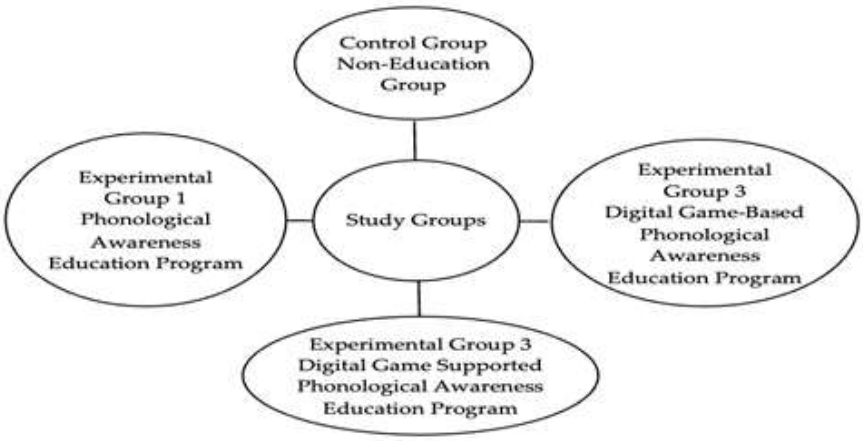
This study examines the effects of different education programs (phonological awareness education program, digital game-supported phonological awareness education program, and digital game-based phonological awareness education program) on the phonological awareness skills of 60-72-month-old children. For this general purpose, Is there a difference between the different education programs applied to the experimental groups in improving the phonological awareness skills of 60-72 children? Is there a difference between the experimental and control groups, in which different educational programs were applied to improve the phonological awareness skills of 60-72 children? Do the education programs applied to the experimental groups increase/improve retention? Answers to these questions were sought.

Method

A quasi-experimental design was used in this research to examine the effects of digital game-based and different educational programs on the phonological awareness skills of 60-72-month-old children. Cohen et al. (2017) state that it is only sometimes possible and challenging to determine the experimental and control groups by random assignment in empirical studies conducted in educational settings. Quasi-experimental designs are studies conducted without random assignment of participants to groups. In this model, the researcher does not assign impartially to the experimental and control groups, and the effect of the applied intervention on the experimental group is examined (Mertens, 2019). Since this research was conducted in an educational institution, a quasi-experimental design was used. Research-ready groups were matched in terms of phonological awareness skills without random assignment. Experimental and control groups were equally unbiased regarding phonological awareness skills. The research consists of three experimental groups and one control group. These groups are shown in Figure 1 below.

Figure 1

Study Groups and Applied Programs of the Research



When examining Figure 1, it can be observed that the research study consists of four groups in the study group: experimental group 1, experimental group 2, experimental group 3, and the control group. The phonological awareness education program was applied to experimental group 1, the phonological awareness education program supported by digital games was applied to experimental group 2, and the digital game-based phonological awareness education program was applied to experimental group 3. The Control group did not receive any specific education and continued with the implementation of the 2013 Preschool Education Program in the classroom. Ministry of National Education [MoNE] (2013) Preschool Education Program is a national preschool education program applied to the control group. However, the program offers flexibility to the teacher regarding which methods will be used to help children achieve the objectives. Therefore, it is at the teacher's discretion which methods the teacher teaching the control group will use for phonological awareness skills. In addition, although there is no technology center in the classroom where the control group receives educational, there is a smart board. However, it was determined that there was no digital game application on the smart board. Pre-test, post-tests, and retention tests were conducted to measure and compare the effectiveness of the programs. Retention testing can be done at least two weeks after educational programs are implemented (Haynie, 1997). In the study, the permanence test was performed three weeks after the last test was applied.

Study Group

In this research, criterion sampling, one of the purposeful sampling methods, was used. Purposive sampling is a sampling technique chosen due to its alignment with the specific purpose of the research and is preferred because of the unique characteristics of the sample elements (Check & Schutt, 2012; Newby, 2010). In this context, it was determined through discussions with school principals and preschool teachers whether educational programs aimed at supporting phonological awareness skills had been previously implemented in nursery schools and preschool classes affiliated with the Ministry of National Education. This approach identified preschool classes where no phonological awareness education program had been participated in, and voluntary participation was sought. The study group of the research consists of a total of 72 children in the 60-72 month age group who received education in preschool classes where a phonological awareness skill-supporting educational program had not been previously implemented within an elementary school affiliated with the Ministry of National Education in Afyonkarahisar city center. Among these are 22 children in experimental group 1, 16 in experimental group 2, 17 in Experimental group 3, and 17 in the Control Group.

In quasi-experimental studies, no precise rule for determining the appropriate sample size exists. However, it is mentioned that conducting a study with groups of 30-40 individuals may provide researchers with various advantages in terms of the generalizability of research results and the availability of robust statistics (Büyüköztürk et al., 2014). The research sample was formed using a cluster sampling technique. Cluster sampling is a method where groups are selected randomly rather than individual individuals. All members of the selected groups have similar characteristics. Any group

with similar characteristics and no inference made is called a cluster (Özen & Gül, 2007). Information about the study group of the research is provided in Table 1.

Table 1

Participant Demographic Information

		Experiment Group 1		Experiment Group 2		Experiment Group 3		Control Group		Total	
		n	%	*n*	%	*n*	%	*n*	%	*n*	%
Age	60-66 Month	8	36.4	6	37.5	11	64.7	10	58.8	35	48.6
	67-72 Month	14	63.6	10	62.5	6	35.3	7	41.2	37	51.4
Gender	Girl	14	63.6	9	56.3	12	70.6	9	52.9	44	61.1
	Boy	8	36.4	7	43.8	5	29.4	8	47.1	28	38.9
Birth Order	Firstborn	4	18.2	4	25	5	29.4	2	11.8	15	20.8
	Middle Child	8	36.4	8	50	6	35.3	3	17.6	25	34.7
	Last Child	10	45.5	4	25	6	35.3	12	70.6	32	44.4
Number of Children in the Family	One	3	13.6	0	0	2	11.8	1	5.9	6	8.3
	Two	14	63.6	11	68.8	10	58.8	8	47.1	43	59.7
	Three	5	22.7	5	31.3	5	29.4	8	47.1	23	31.9
Attending a Preschool Institution	0-6	1	4.5	2	12.5	2	11.8	1	5.9	6	8.3
	7-12	13	59.1	13	81.3	8	47.1	15	88.2	49	68.1
	13-18	4	18.2	0	0	3	17.6	1	5.9	8	11.1
	19-24	2	9.1	1	6.3	2	11.8	0	0	5	6.9
	25 and over	2	9.1	0	0	2	11.8	0	0	4	5.6

When Table 1 is examined, it is seen that there are 22 children in experimental group 1, 16 children in experimental group 2, 17 children in experimental group 3, and 17 children in the control group. When the ages of the children are examined, in experimental group 1, 60-66 month-old children are 8 (36.4%), 67-72 month-old children are 14 (63.6%), in experimental group 2, 60-66 months old 6 (37.5), 10 (62.5%) if 67-72 months old, 11 (64.7%) 60-66 months old, 6 (35.3%) 67-72 months old, 60-66 months old when the control group is examined It is seen that the children are in the age range of 10 (58.8%) and 7 (41.2%) between 67-72 months. When the children are examined in terms of gender, it is seen that the number of girls is 44 (61.1%) and the number of boys is 28 (38.9%). When examined in order of birth, the children participating in the study were the last child at most, 32 (44.4%) and the least the firstborn was 15 (20.8%). Considering the number of children in the family, there are at most 2 (63.6%) and at least 6 (8.3%) single children in the families of the children participating in the study. Considering the pre-school education status of the children participating in the research, it is seen that 49 (68.1%) children attend school between 7 and 12 months and continue for at least 25 months and above (5.6%).

Data Collection Tools

In the study, the Early Literacy Assessment Tool (ELST) and the Early Literacy Test for Kindergarten Children (ELTKC) were used to evaluate the phonological awareness skills of the experimental and control groups.

Early Literacy Skills Assessment Tool (ELST)

The measure used to examine early literacy skills in preschool children was developed by Karaman and Aytar (2016). It applies to children between the ages of 48 and 77 months. The scale consists of 5 subtests: phonological awareness, story narration, matching pictures, writing awareness, and pre-writing skills. This study used the phonological awareness subtest as the research objectiveIn the phonological awareness skills subtest, there are five factors, which include matching rhyming words (9 items), matching words with the same initial sound (6 items), identifying initial sounds (21 items), blending sounds (7 items), and segmenting syllables and sounds (10 items). The reliability of the Phonological Awareness Skills Assessment subtest is indicated by a KR20 reliability coefficient of .91, and the test-retest reliability value is .92. The test's scoring is as follows: incorrect responses receive 0 points, correct responses receive 1 point. In the case of no response, the question item is repeated three times, and if there is still no response, 0 points are given. The test is administered individually and takes approximately 45-60 minutes. The evaluation of the test involves calculating the average score obtained by the children in the group, and then each child's score is assessed concerning the average score.

Early Literacy Test for Kindergarten Children (ELTKC)

The measurement tool used to assess early literacy skills in preschool children was developed by Kargın et al. (2017). It consists of subtests for expressive language, fluent language, functional knowledge, general naming, listening comprehension, phonological awareness, and letter knowledge. In this study, the phonological awareness subtest of the test was used. The phonological awareness subtest also comprises eight subtests, which include segmenting sentences into words, syllable segmentation, rhyme awareness, matching initial sounds, matching final sounds, syllable blending, deleting final sounds, and deleting initial sounds. Each subtest contains four items. The overall reliability coefficient of the scale is .94. The Spearman-Brown two-half reliability value is .79. The reliability of the phonological awareness subtest is indicated by a KR20 value of .87, the test-retest reliability coefficient of .70, and a Spearman-Brown two-half reliability value of .67. The scale can be administered to children in the age group of 60-72 months. Scoring for the test involves assigning 0 points for incorrect answers and 1 point for correct answers. Assessment is conducted based on the cutoff point, and the cutoff point for the phonological awareness subtest is .50.

When examining the assessment tools used in the research, it is observed that the sub-factors within phonological awareness tests differ. Different assessment tools have been used to comprehensively investigate the effectiveness of the applied educational programs in the research.

Analysis of Data

Since the number of participants in the experimental and control groups is less than thirty, non-parametric analysis methods can be employed for data analysis (Büyüköztürk, 2013). Consequently, the Mann-Whitney U test has been used to compare the pre-test data. Additionally, the Wilcoxon Signed-Rank Test has been utilized to compare the pre-test to the post-test and the post-test to retention test scores.

Preparation and Implementation of Education Programs

Within the scope of the research, a phonological awareness education program, a digital game-based phonological awareness education program, and a digital game-supported phonological awareness education program were developed. During the preparation process of the program, the aim was to create a comprehensive program focusing on phonological awareness skills. Therefore, a review of the relevant literature and scales related to phonological awareness skills was conducted. As a result of these reviews, decisions were made regarding which phonological awareness skills would be covered in the program. It was then compared with the achievements and indicators for phonological awareness in the currently implemented Preschool Education Program for 36-72 72-month-old children (MoNE, 2013). As a result of the comparisons, it was seen that there were differences between the phonological awareness gains in the program and the phonological awareness skills determined for the research. Due to these differences, skills such as word awareness, syllable segmentation, determining the number of syllables in a word, syllable deletion, blending syllables, deleting the first/last sound in a word, and sound blending were added to the outcomes and indicators of the educational programs implemented in the research. After determining the outcomes and indicators, the educational programs were prepared. The prepared educational programs were child-centered, spiral, gamified, actively participatory, and interactive and aimed to transition from the known to the unknown, avoiding didactic approaches. Furthermore, the digital game was prepared following Gestalt theory, Gagne's instructional conditions model, and cognitive load theory. After the educational programs were developed, they were sent to three academics specializing in program development for evaluation. The academics stated that the programs were appropriate regarding outcomes, indicators, the learning process, assessment, and materials. Subsequently, the programs were implemented with the experimental groups. The first researcher carried out the implementation of the educational programs prepared for the experimental groups. In the spring semester of the 2021-2022 academic year, practices were carried out in the classrooms where children were educated on Tuesdays and Wednesdays of the week. The researcher attended the children's education for three days so that the children could get to know the researcher and establish interaction between them. Then, the pre-test and implementation phase started. A phonological awareness educational program was applied to the Experiment 1 group. In this educational program, various activities were carried out using different methods in Turkish, drama, mathematics, music, games and pre-literacy preparation activity types and activity types. For example, methods and techniques such as shared reading and drama were used in the Turkish activity, and problem solving in the music activity. Natural objects, toys or object images, game cards, puppets and similar materials were used in these activities. The activities were carried out in the form of individual, large and small studies. Care was taken to ensure a diverse array of materials in the activities, providing with the opportunity to choose materials, reflect on their own thoughts, and have their preferences implemented. For example, by allowing children to create poems as a group and stories individually, add different words to songs, use different materials in the classroom, and add rules to games (which will not affect the acquisition of phonological awareness skills), the education process was designed to be flexible, ensuring active participation from the children. Additionally, a listening and writing

center was established in the educational environment for both Experimental group 1 and Experimental group 2, and adjustments were made in other centers according to the activities included in the program. The centers were updated with different materials every week. Experimental group 1 received the phonological awareness education program for ten weeks, with two weekly sessions, each lasting approximately 30 minutes.

Experimental group 3 received the digital game-based phonological awareness education program, administered for ten weeks with two sessions per week. The duration of each session ranged from approximately 20-25 minutes. The digital game reflects phonological awareness skills. First of all, sentences, words, and instructions to be presented to children reflecting each phonological awareness sub-skill were determined. In determining the sentences, words and instructions, Memoğlu-Süleymanoğlu's (2014a) Turkish Frequency Dictionary and Turkish Reverse Frequency Dictionary (2014b) studies were utilized (Gökmen, 2007; Görgün, 2020; Keklik, 2010, 2011; Şahin Kamışlı et al., 2015; Savaş & Turan, 2011; Topbaş, 1996; Topbaş, 2006). Then, it was sent to academics working in the field of curriculum to check the appropriateness of words, sentences, and instructions. In line with the feedback from the academics, words and sentences that were not suitable for children were removed, and the instructions were organized. For example, the word kaleidoscope was removed, considering that it is a word that children rarely encounter in their daily lives. Afterwards, a meeting was held with the Sekizdesekiz company, and a game was prepared using words, sentences, and instructions. Visuals suitable for the words, sentences and instructions were prepared and vocalized by Firma. This process continued with an exchange of views between the researchers and the company. For example, when it was thought that children could not make sense of the prepared visual, the visual was drawn differently. Or the necessary ideas were exchanged to ensure that the vocalization attracted children's attention and that the word or sentence was clearly understood. In this process, the visuals and vocalizations were transferred to the game application Sekizdesekiz used. The company's game application can be installed on smart boards, computers, and tablets. In the research, the application was installed on smart boards. In the process of playing digital games, the study was conducted as a large group so that each child could play the game at the same time. The tool for this was the remote controllers prepared by the company for large group work. A remote control was defined for each child, and children used their own remote controls during the application process. After the game application is opened, each child defines his/her own remote control. Next comes the interface of the game (visuals, the spelling of words, background, etc.), followed by the instructions for the sentences, words, and voiceover. This allows children to listen and see, appealing to multiple senses. After listening to the instructions and words, children tap the remote control button to give the correct answer.

Experimental group 2 received the digital game-supported phonological awareness education program. For this group, the phonological awareness education program was applied, which included activities using natural objects, toys or visuals of objects, game cards, puppets, and similar materials (Turkish, drama, mathematics, music, play, and pre-literacy preparation). After each application, a digital game explicitly prepared for that skill was played. For example, after participating in an

activity related to rhyme awareness, the children played a digital game related to the rhyme. The program applied to Experimental group 2 lasted for ten weeks, with two sessions per week, each lasting approximately 45-50 minutes.

The educational programs prepared for the experimental groups were structured sequentially according to the phonological awareness dimensions specified in the literature and included in the scales. For example, activities and digital games start with sentence awareness and then continue with word awareness. After spelling the words and determining the number of syllables in the word, studies are carried out on rhyme awareness. Additionally, every three weeks, different materials, activities, and words related to phonological awareness skills that had been previously covered were introduced. This process also applies to the digital game. In the digital game, the activities were arranged sequentially based on phonological awareness skills, and every three weeks, review games with different words and visuals were played.

Ethical Procedures

An application was made to the Ethics Committee of the University to conduct the research on 06.12.2021. The application to the ethics committee included a file containing information about the research's objectives, the measurement tools to be used in the research, and details about the researchers. The ethics committee members reviewed the file, and approval for the research was granted on 14.01.2022. The ethics committee approval bears the date and file number 14.01.2022-E.264744.

Results

This section presents the findings related to the pre-test, post-test, and retention test based on the ELTKC and ELST assessment tools for the experimental groups.

Comparison of Pre-test and Post-test Total Scores of Experimental Groups

In the research, to compare the effectiveness of the educational programs applied to the experimental groups on children's phonological awareness skills, a Kruskal-Wallis H test was conducted to determine whether there was a significant difference in the total scores of the pre-test and post-test based on the ELTKC assessment tool. The results are presented in Table 2.

Table 2

Kruskal-Wallis H Test Results for Pre-Test and Post-Test Total Scores of Experimental Groups based on ELTKC

	Experimental Groups	*n*	Mean Ranks	x^2	*df*	*p**	Significant Difference
ELTKC Pre-Test Total	Experimental Group 1	22	42.18	5.495	3	.139	-
	Experimental Group 2	16	40.84				
	Experimental Group 3	17	28.09				
	Control Group	17	33.47				
ELTKC Post-Test Total	Experimental Group 1	22	48.02	32.994	3	.000*	4<1 4<2 4<3
	Experimental Group 2	16	46.50				
	Experimental Group 3	17	36.21				
	Control Group	17	12.47				
	Total	72					

*p<.05

When examining the ELTKC post-test total scores, a notable difference is evident among experimental groups 1, 2, 3 when compared to the control group without education (x^2 (df=3, n=72)=5.495, p=.000, p>.05). However, when the relationship between experimental group 1, experimental group 2, and experimental group 3 is examined, no significant difference is observed among the groups. Therefore, the educational programs applied to experimental group 1, experimental group 2, and experimental group 3 were effective compared to the control group.

The ELTKC pre-test and post-test scores of experimental group 1, experimental group 2, experimental group 3, and the control group within the experimental group were examined in terms of their means relative to the cutoff point of ELTKC. The findings are presented in Table 3.

Tablo 3

Grouping of Pre-Test and Post-Test Scores According to ELTKC Cutoff Point

	Pre-test				Post-test			
	Below the Cutoff Point		Above the Cutoff Point		Below the Cutoff Point		Above the Cutoff Point	
Groups	*n*	%	*n*	%	*n*	%	*n*	%
Experimental Group 1	18	81.8	4	18.2	1	4.5	21	95.5
Experimental Group 2	12	75.0	4	25.0	2	12.5	14	87.5
Experimental Group 3	17	100.0	0	0.0	1	5.9	16	94.1
Control Group	16	94.1	1	5.9	10	58.8	7	41.2

When examining the values in Table 3, it can be observed that in Group 1 of the experimental group, there were 18 children below the cutoff point and four children above the cutoff point in the ELTKC pre-test. After implementing the phonological

awareness education program in Group 1, in the ELTKC post-test, 21 children in Group 1 were above the cutoff point, and one child was below the cutoff point. Looking at Group 2, according to the ELTKC pre-test results, there were 12 children below the cutoff point and four children above the cutoff point in Group 2. After applying the digital game-supported phonological awareness education program to Group 2, in the ELTKC post-test, there were two children below and 14 children above the cutoff point in Group 2. Analyzing the data for Group 3, all the children in Group 3 are below the cutoff point in the ELTKC pre-test. After playing the digital game, when the ELTKC post-test results were examined, there was one child below the cutoff point and 16 children above the cutoff point in Group 3. Finally, looking at the control group, according to the ELTKC pre-test results, there are 16 children below the cutoff point and one child above the cutoff point. When examining the post-test results of the control group, there were ten children below and seven above the cutoff point.

In the study, to compare the effects of the education programs applied to the experimental groups on children's phonological awareness skills, the total mean scores of the groups were calculated using the Kruskal-Wallis H test to determine if there was a significant difference in the ELST pre-test and post-test total scores. The results are presented in Table 4.

Table 4

Results of Kruskal-Wallis H Test for ELST Pre-Test and Post-Test Total Scores Between Experimental Groups

	Experimental Groups	*n*	Mean Ranks	x^2	*df*	*p**	Significant Difference
ELST Pre- Test Total	Exp. Group 1	22	40.41	2.724	3	.436	-
	Exp. Group 2	16	40.09				
	Exp. Group 3	17	33.29				
	Control Group	17	31.26				
ELST Post- Test Total	Exp. Group 1	22	50.50	38.017	3	.000*	3<1 4<1 3<2 4<2 4<3
	Exp. Group 2	16	48.19				
	Exp. Group 3	17	31.18				
	Control Group	17	12.71				
	Total	72					

*p<.05

Table 4 reveals that there was no significant difference among the experimental groups in terms of ELST pre-test total scores (x^2 (*df*=3, *n*=72)=2.724, *p*=.436, *p*>.05). Hence, the experimental groups had similar phonological awareness skills at the beginning. However, when examining the ELST post-test total scores, it can be observed that there was a significant difference between experimental groups 1, 2, and 3 compared to the control group, indicating an increase in phonological awareness skills in the experimental groups after the implementation of the education program. Moreover, a significant difference was observed when comparing the experimental

groups among themselves. This significant difference favors experimental group 1 when compared to experimental group 3. There was no significant difference between experimental group 1 and experimental group 2. Nevertheless, there was a more significant increase in scores for experimental group 1 compared to experimental group 2 after implementing the education programs. Additionally, when comparing experimental group 2 to experimental group 3, a significant difference was observed in favor of experimental group 2.

The Examination of the Pre-Test Scores of the Experimental Groups Based on Sub-dimensions

In the study, the Kruskal Wallis H test was conducted to determine whether there was a significant difference in the ELTKC pre-test scores of the experimental groups and the control group regarding phonological awareness skills, and the results are presented in Table 5.

Table 5

Results of Kruskal Wallis H Test for Group Differences in Pre-Test Score Averages by ELTKC Sub-dimensions

	Experimental Groups	*n*	Mean Rank	x^2	*df*	*p**
Rhyme Awareness	Experimental Group 1	22	36.68	4.019	3	.259
	Experimental Group 2	16	43.09			
	Experimental Group 3	17	37.26			
	Control Group	17	29.29			
Matching by Initial Sound	Experimental Group 1	22	38.95	2.713	3	.438
	Experimental Group 2	16	36.22			
	Experimental Group 3	17	30.00			
	Control Group	17	40.09			
Matching by Final Sound	Experimental Group 1	22	35.93	3.164	3	.367
	Experimental Group 2	16	37.66			
	Experimental Group 3	17	30.50			
	Control Group	17	42.15			
Segmenting Sentences into Words	Experimental Group 1	22	43.75	5.653	3	.130
	Experimental Group 2	16	34.56			
	Experimental Group 3	17	29.32			
	Control Group	17	36.12			
Segmenting Words into Syllables	Experimental Group 1	22	40.09	6.834	3	.077
	Experimental Group 2	16	34.56			
	Experimental Group 3	17	29.32			
	Control Group	17	36.12			
Blending Syllables	Experimental Group 1	22	43.00	3.986	3	.263
	Experimental Group 2	16	35.19			
	Experimental Group 3	17	32.32			

Delete the Initial Sound in Words	Control Group	17	33.50	4.766	3	.190
	Experimental Group 1	22	38.95			
	Experimental Group 2	16	38.44			
	Experimental Group 3	17	34.00			
Delete the Final Sound in Words	Control Group	17	34.00	4.176	3	.243
	Experimental Group 1	22	37.80			
	Experimental Group 2	16	40.03			
	Experimental Group 3	17	35.00			
	Control Group	17	33.00			
	Total	72				

*p<.05

When examining Table 5, it can be observed that there was no significant difference in the pre-test scores and total scores of the ELTKC sub-dimensions among the experimental groups and the control group (x^2 (df=3, n=72)=5.495, p>.05). Therefore, the phonological awareness skill levels of the control and experimental groups were similar concerning ELTKC sub-dimensions.

In the study, a Kruskal Wallis H test was conducted to determine whether there was a significant difference in the ELST pre-test scores among the experimental and control groups, and the results are shown in Table 6.

Table 6

Results of Kruskal Wallis H Test for Mean Pre-Test Scores by ELST Sub-Dimensions Among Groups

	Experimental Groups	*n*	Mean Rank	x^2	*df*	*p*
Matching Words Beginning with the Same Sound	Experimental Group 1	22	35.82	1.260	3	.739
	Experimental Group 2	16	41.31			
	Experimental Group 3	17	34.62			
	Control Group	17	34.74			
Matching Rhyming Words	Experimental Group 1	22	44.34	6.003	3	.111
	Experimental Group 2	16	35.34			
	Experimental Group 3	17	28.91			
	Control Group	17	35.03			
Finding the Beginning Sound of the Given Word	Experimental Group 1	22	37.64	3.487	3	.322
	Experimental Group 2	16	35.81			
	Experimental Group 3	17	40.29			
	Control Group	17	31.88			
Producing Words That Begin with a Stimulus Sound	Experimental Group 1	22	37.36	.757	3	.860
	Experimental Group 2	16	36.44			
	Experimental Group 3	17	33.09			
	Control Group	17	38.85			

Producing Words That Start With the Same Sound	Experimental Group 1	22	38.30	1.764	3	.623
	Experimental Group 2	16	40.38			
	Experimental Group 3	17	34.82			
	Control Group	17	32.21			
Delete Syllables and Sounds	Experimental Group 1	22	37.64	1.673	3	.643
	Experimental Group 2	16	41.06			
	Experimental Group 3	17	34.76			
	Control Group	17	32.47			
Merge Sounds	Experimental Group 1	22	38.41	4.439	3	.218
	Experimental Group 2	16	40.66			
	Experimental Group 3	17	27.74			
	Control Group	17	38.88			
	Total	72				

*p<.05

When Table 6 is examined, it can be observed that there was no significant difference in the mean ELST pre-test scores among the experimental groups and the control group (x^2 (df=3, n=72)=2.724, p>.05). Therefore, the phonological awareness skill levels of the control and experimental groups were similar according to ELST sub-dimensions.

Examination of Experimental Groups' Final Test Scores by Sub-Dimensions

In the study, in order to determine whether there was a significant difference in the phonological awareness skills of the experimental groups and the control group based on the ELTKC post-test scores, a Kruskal-Wallis H test was conducted, and the results are presented in Table 7.

Table 7

Results of Kruskal-Wallis H Test for Mean Scores on ELTKC Sub-Dimensions in Terms of Post-Test Scores by Groups

	Experimental Groups	n	Mean Rank	x^2	df	p	Significant Difference
Rhyme Awareness	Experimental Group 1	22	46.82	38.329	3	.000*	4<1 4<2 4<3
	Experimental Group 2	16	43.69				
	Experimental Group 3	17	40.65				
	Control Group	17	12.24				
Matching by Initial Sound	Experimental Group 1	22	42.18	28.700	3	.000*	4<1 3<2 4<2 4<3
	Experimental Group 2	16	50.88				
	Experimental Group 3	17	36.29				
	Control Group	17	15.82				
Matching by Final Sound	Experimental Group 1	22	43.86	33.093	3	.000*	4<1 3<2
	Experimental Group 2	16	49.69				

	Experimental Group 3	17	37.88				4<2
	Control Group	17	13.18				4<3
	Experimental Group 1	22	45.64				
Segmenting Sentences into Words	Experimental Group 2	16	39.84	36.541	3	.000*	4<1
	Experimental Group 3	17	42.97				4<2
	Control Group	17	15.06				4<3
	Experimental Group 1	22	41.41				
Segmenting Words into Syllables	Experimental Group 2	16	38.22	18.069	3	.000*	4<1
	Experimental Group 3	17	40.94				4<2
	Control Group	17	24.09				4<3
	Experimental Group 1	22	39.50				
Blending Syllables	Experimental Group 2	16	34.88	4.334	3	.228	4<1
	Experimental Group 3	17	37.41				
	Control Group	17	33.24				
	Experimental Group 1	22	45.98				
Delete the Initial Sound in Words	Experimental Group 2	16	38.16	10.325	3	.016*	2<1
	Experimental Group 3	17	30.06				4<1
	Control Group	17	29.12				
	Experimental Group 1	22	45.09				
Delete the Final Sound in Words	Experimental Group 2	16	43.91	22.088	3	.000*	4<1
	Experimental Group 3	17	37.91				4<2
	Control Group	17	17.00				4<3
	Total	72					

*p<.05

Table 7 indicates that there was a significant difference in the mean scores of the experimental groups and the control group in all sub-dimensions of ELTKC post-test scores except for the "syllable merging" sub-dimension (x^2 (n=72; df=3)=32.994, p<.05). According to the results of the Mann-Whitney U test conducted to determine which groups differ, it has been found that the experimental groups scored significantly higher than the control group in all areas except for the "syllable merging" and "producing words that start with the same sound" sub-dimensions. Specifically, only experimental group 1 had higher mean scores than the control group in all areas, and additionally, experimental group 2 had significantly higher mean scores than experimental group 3 in the "matching words beginning with the same sound" and "matching rhyming words" sub-dimensions.

In the research, a Kruskal Wallis H test was conducted to determine whether there was a significant difference in ELST post-test scores between the experimental and control groups, and the results are shown in Table 8.

Table 8

Results of the Kruskal Wallis H Test for Inter-Group Comparison of Mean Scores of ELST Sub-dimensions at Post-Test

	Experimental Groups	*n*	Mean Rank	x^2	*df*	*p*	Significant Difference
Matching Words Beginning with the Same Sound	Experimental Group 1	22	48.59	31.086	3	.000	3<1 4<1 4<2 4<3
	Experimental Group 2	16	45.50				
	Experimental Group 3	17	34.50				
	Control Group	17	14.38				
Matching Rhyming Words	Experimental Group 1	22	49.70	37.521	3	.000	3<1 4<1 3<2 4<2 4<3
	Experimental Group 2	16	48.41				
	Experimental Group 3	17	31.32				
	Control Group	17	13.38				
Finding the Beginning Sound of the Given Word	Experimental Group 1	22	43.18	19.419	3	.000	4<1 4<2 4<3
	Experimental Group 2	16	39.94				
	Experimental Group 3	17	38.85				
	Control Group	17	22.26				
Producing Words That Begin with a Stimulus Sound	Experimental Group 1	22	49.09	32.852	3	.000	3<1 4<1 3<2 4<2 4<3
	Experimental Group 2	16	45.66				
	Experimental Group 3	17	33.82				
	Control Group	17	14.26				
Producing Words That Start With the Same Sound	Experimental Group 1	22	44.82	27.514	3	.000	4<1 4<2 4<3
	Experimental Group 2	16	45.56				
	Experimental Group 3	17	39.50				
	Control Group	17	14.21				
Delete Syllables and Sounds	Experimental Group 1	22	48.57	33.211	3	.000	3<1 4<1 4<2 4<3
	Experimental Group 2	16	46.34				
	Experimental Group 3	17	35.12				
	Control Group	17	13.00				
Merge Sounds	Experimental Group 1	22	45.70	29.854	3	.000	4<1 4<2 4<3
	Experimental Group 2	16	47.72				
	Experimental Group 3	17	36.79				
	Control Group	17	13.74				
ELST Post-Test Total	Experimental Group 1	22	50.50	38.017	3	.000	3<1 4<1 3<2 4<2 4<3
	Experimental Group 2	16	48.19				
	Experimental Group 3	17	31.18				
	Control Group	17	12.71				
	Total	72					

*p<.05

Table 8 shows that there was a significant difference in the mean scores of ELST post-test sub-dimensions among the experimental groups and the control group (x^2 (n=72, df=3)=38.017, p<.05). According to the results of the Mann-Whitney U test conducted to determine which groups differ, it is observed that the experimental groups scored significantly higher than the control group in all dimensions. Additionally, in the sub-dimensions of matching words beginning with the same sound, syllables, and sounds dropping, both experimental groups 1 and 2 scored higher than experimental group 3, and in the sub-dimensions of matching rhyming words, noticing the initial sounds of words, and the total score, both experimental groups 1 and 2 scored higher than experimental group 3.

Examination of the Retention Test Scores of the Experimental Groups

The Wilcoxon Signed-Rank Test was conducted to examine the significant differences between the final test and retention test scores of Experimental group 1, Experimental group 2, and Experimental group 3 in terms of their auditory awareness skills in ELTKC and ELST. The results are displayed in Table 9 and Table 10.

The Wilcoxon Signed-Rank Test examined the significant differences between the final test and retention test scores of Experimental group 1, Experimental group 2, and Experimental group 3 regarding their auditory awareness skills in ELTKC and ELST. The results are displayed in Table 9 and Table 10. It was determined that the scores in the sub-dimensions of matching to the initial sound (Positive Rank Sum ST=28.00) and matching to the final sound (Positive Rank Sum ST=66.00) increased. At the same time, there was a decrease in the dimension of omitting the initial sound of words (Negative Rank Sum ST=70.00). It can be observed that the mean scores of the final test and retention test in Experimental group 1 are similar. A significant difference was observed between the ELTKC final test and retention test scores in Experimental group 2 (n=16; Z=-2.289; p<.05). However, when examining the sub-dimensions of ELTKC, no significant differences were found in any of the sub-dimensions (p<.05). Upon closer examination of the sub-dimensions of ELTKC, a decrease was observed in the sub-dimension matching the initial sound (Negative Rank Sum ST=10.00). At the same time, an increase was noted in the positive rank sums in other sub-dimensions and the total positive rank sum (Total Positive Rank Sum ST=68.00). Despite the observed decreases and increases in the sub-dimensions of ELTKC, these values did not show a significant decrease or increase. There is a difference between the final test and retention test scores, with an increase in Experimental group 2's ELTKC retention test score. On the other hand, there was no significant difference found between the ELTKC final test and retention test scores in Experimental group 3 (n=17; Z=-.595; p<.05). When examined according to the sub-dimensions of ELTKC, no significant differences were found in any of the sub-dimensions (p<.05). When looking at the total rank sum value of ELTKC, it can be observed that there is a slight increase in Experimental group 3 overall (Total Positive Rank Sum ST=153.00). However, it can be concluded that this increase is not statistically significant. There is also a slight increase in the ELTKC retention test results for Experimental group 3, but according to the analysis conducted, this increase does not reflect a significant difference.

When examining the findings in Table 10, it is evident that there is a significant difference between the ELST final test and retention test scores in Experimental group 1

(z=-2.112; p<.05). When looking at the sub-dimensions of ELST, a significant difference is observed in the sub-dimension of matching words that start with the same sound (p<.05), and there is a decrease in the correct answers given in this sub-dimension by Experimental group 1 (Negative Rank Sum ST=66.00). There is no significant difference found in the sub-dimensions other than matching words that start with the same sound (p<.05). However, it is observed that there is a decrease in the answers provided by Experimental group 1 (Total Negative Rank Sum ST=108.50). There is a difference between the final test and retention test scores, with a decrease in the ELST retention test score of Experimental group 1 compared to the final test score. Although this difference is slight, it is statistically significant, as seen in Table 10. In contrast, no significant difference was observed in the ELST final test and retention test scores in Experimental group 2 (n=16; Z=-.361; p<.05).

When looking at the sub-dimensions of ELST, it was found that there is no significant difference in any of the sub-dimensions of ELST (p<.05). However, when examining the sub-dimensions of ELST, there is a decrease in the sub-dimension of omitting syllables and sounds (Negative Rank Sum ST=57.00), and there is also a decrease in the total rank sum of ELST. There is a slight decrease in the average score of the retention test for experimental group 2 compared to the final test score. However, this decrease reflects a slight difference. A significant difference was found between the ELST final test and retention test scores in experimental group 3 (n=17; Z=-2.080; p<.05). When examined according to the sub-dimensions of ELST, there is a significant difference in the sub-dimension of matching words that start with the same sound. However, there is no significant difference in the other sub-dimensions. When examining ELST's total rank sum value, there is an increase in experimental group 3. When looking at the rank sum values of the sub-dimensions, there is an increase in the sub-dimensions of matching words that start with the same sound (Positive Rank Sum ST=7.00) and omitting syllables and sounds (Positive Rank Sum ST=45). There is a slight decrease in the sub-dimension of matching rhyming words (Negative Rank Sum ST=33.00). Experimental group 3 shows an increase in the average score of the ELST retention test compared to the final test score, which results in a significant difference.

Table 9

Groups ELTKC Retention Test Results

	Experimental Groups	Negative Rank (n)	Negative Rank (R.A)	Negative Rank (R.S)	Positive Rank (n)	Positive Rank (R.A)	Positive Rank (S.T)	Equal (n)	z	p*
Rhyme Awareness	Exp. Group 1	0	.00	.00	3	2.00	6.00	19	-1.732	.083
	Exp. Group 2	0	.00	.00	3	2.00	6.00	13	-1.732	.083
	Exp. Group 3	4	4.00	16.00	4	5.00	20.00	9	-.302	.763
Matching by Initial Sound	Exp. Group 1	0	.00	.00	7	4.00	28.00	15	-2.530	.011
	Exp. Group 2	3	3.33	10.00	2	2.50	5.00	11	-.707	.480
	Exp. Group 3	2	4.00	8.00	5	4.00	20.00	10	-1.134	.257
Matching by Final Sound	Exp. Group 1	0	.00	.00	11	6.00	66.00	11	-3.317	.001*
	Exp. Group 2	2	4.50	9.00	4	3.00	12.00	10	-.333	.739
	Exp. Group 3	2	4.00	8.00	5	4.00	20.00	10	-1.134	.257

Segmenting Sentences into Words	Exp. Group 1	2	2.00	4.00	1	2.00	2.00	19	-.577	.564
	Exp. Group 2	0	.00	.00	3	2.00	6.00	13	-1.604	.109
	Exp. Group 3	1	1.50	1.50	2	2.25	4.50	14	-.816	.414
Segmenting Words into Syllables	Exp. Group 1	2	1.50	3.00	0	.00	.00	20	-1.414	.157
	Exp. Group 2	0	.00	.00	2	1.50	3.00	14	-1.342	.180
	Exp. Group 3	1	1.00	1.00	0	.00	.00	16	-1.000	.317
Blending Syllables	Exp. Group 1	0	.00	.00	0	.00	.00	22	.000	1.000
	Exp. Group 2	0	.00	.00	1	1.00	1.00	15	-1.000	.317
	Exp. Group 3	1	2.00	2.00	1	1.00	1.00	15	-.447	.655
Delete the Initial Sound in Words	Exp. Group 1	11	6.36	70.00	1	8.00	8.00	10	-2.470	.014
	Exp. Group 2	0	.00	.00	4	2.50	10.00	12	-1.890	.059
	Exp. Group 3	4	3.13	12.50	1	2.50	2.50	12	-1.414	.157
Delete the Final Sound in Words	Exp. Group 1	3	5.00	15.00	8	6.38	51.00	11	-1.706	.088
	Exp. Group 2	2	6.00	12.00	8	5.38	43.00	6	-1.628	.103
	Exp. Group 3	1	2.00	2.00	3	2.67	8.00	13	-1.134	.257
Total	Exp. Group 1	6	7.42	44.50	9	8.39	75.50	7	-.887	.375
	Exp. Group 2	1	10.00	10.00	11	6.18	68.00	4	-2.289	.022*
	Exp. Group 3	4	6.63	26.50	7	5.64	39.50	6	-.595	.552

Table 10

Experimental Groups ELST Retention Test Results

	Experimental Groups	Negative Rank (n)	Negative Rank (R.A)	Negative Rank (R.S)	Positive Rank (n)	Positive Rank (R.A)	Positive Rank (S.T)	Equal (n)	z	p*
Matching Words Beginning with the Same Sound	Exp. Group 1	10	6.60	66.00	2	6.00	12.00	10	-2.324	.020*
	Exp. Group 2	5	3.50	17.50	2	5.25	10.50	9	-.632	.527
	Exp. Group 3	2	5.50	11.00	10	6.70	67.00	5	-2.352	.019
Matching Rhyming Words	Exp. Group 1	6	5.33	32.00	4	5.75	23.00	12	-.471	.638
	Exp. Group 2	4	3.50	14.00	3	4.67	14.00	9	.000	1.000
	Exp. Group 3	6	5.50	33.00	5	6.60	33.00	6	.000	1.000
Finding the Beginning Sound of the Given Word	Exp. Group 1	1	2.00	2.00	1	1.00	1.00	20	-.447	.655
	Exp. Group 2	0	.00	.00	1	1.00	1.00	15	-1.000	.317
	Exp. Group 3	3	2.00	6.00	0	.00	.00	14	-1.633	.102
Producing Words That Begin with a Stimulus Sound	Exp. Group 1	8	5.13	41.00	2	7.00	14.00	12	1.430	.153
	Exp. Group 2	2	2.00	4.00	2	3.00	6.00	12	-.378	.705
	Exp. Group 3	4	4.63	18.50	5	5.30	26.50	8	-.491	.623
Producing Words That Start With the Same Sound	Exp. Group 1	5	4.50	22.50	7	7.93	55.50	10	-1.344	.179
	Exp. Group 2	3	4.00	12.00	3	3.00	9.00	10	-.333	.739
	Exp. Group 3	5	5.50	27.50	5	5.50	27.50	7	.000	1.000
Delete Syllables and Sounds	Exp. Group 1	12	8.63	103.50	4	8.13	32.50	6	-1.866	.062
	Exp. Group 2	10	5.70	57.00	2	10.50	21.00	4	-1.446	.148
	Exp. Group 3	6	7.67	46.00	7	6.43	45.00	4	-.036	.971

Merge Sounds	Exp. Group 1	6	6.17	37.00	8	8.50	68.00	8	-1.008	.313
	Exp. Group 2	5	5.40	27.00	4	4.50	18.00	7	-.577	.564
	Exp. Group 3	1	5.50	5.50	5	3.10	15.50	11	-1.081	.279
ELST Post-Test Total	Exp. Group 1	13	8.35	108.50	3	9.17	27.50	6	-2.112	035*
	Exp. Group 2	7	6.21	43.50	5	6.90	34.50	4	-.361	.718
	Exp. Group 3	2	10.00	20.00	12	7.08	85.00	3	-2.080	038*

*p<.05

Discussion and Conclusion

Before proceeding to the results and discussion, it is recommended to revisit Figure 1 to understand better which educational programs were applied to the experimental groups.

When examining the total scores for ELTKC pre-tests and post-tests, the total scores for ELST pre-tests and post-tests, the number of children who scored below the cutoff point based on pre-test results, and also the pre-test data for ELTKC sub-dimensions six and the pre-test data for ELST sub-dimensions, it was determined that there were no significant differences among experimental group 1, experimental group 2, experimental group 3, and the control groups, both in total and sub-dimensions. The absence of significant differences may indicate that the levels of auditory awareness skills were similar among these groups. This situation may be suitable for evaluating and comparing the effect of the educational programs applied to the experimental groups on auditory awareness skills.

Comparison Among the Experimental Groups

Upon examination, it is determined that, based on the ELTKC final test total scores, the applied educational programs positively affected the auditory awareness skills of the experimental groups. However, based on the final test total scores, there is no significant difference between experimental group 1, experimental group 2, and experimental group 3. Therefore, a comparison regarding the applied educational programs cannot be made based on ELTKC pre-test and post-test total scores. According to this data, the educational programs applied to the experimental groups had similar effects.

When the number of children in experimental group 1 who scored above and below the cutoff score on the ELTKC final test was examined, it was found that 21 children in experimental group 1 scored above the cutoff score. In contrast, one child scored below the cutoff score. In experimental group 3, 16 children scored above the cutoff score, while one scored below the cutoff score. In experimental group 2, 14 children scored above the cutoff score, while two children scored below the cutoff score. In the control group where the educational program was not implemented, it was observed that seven children scored above the cutoff score, while ten children scored below the cutoff score. The applied educational programs contributed to the phonological awareness skills of the children in the experimental group. Additionally, it is observed that the number of children scoring below the cutoff point has decreased in the control group, where the educational program still needs to be implemented. The decrease in the number of children scoring below the cutoff point in the control group

can be attributed to the fact that the control group continued with the 2013 Preschool Education Program, which includes rhyme awareness and phonological awareness indicators in the language development domain. The presence of these skills in the education program received by the children in the control group may be the source of the meaningful difference.

When the ELST final test scores were examined, it was concluded that there was a significant difference between the experimental groups and the control groups. The educational programs applied to the experimental groups positively affected the children's phonological awareness skills. When the final test total scores of experimental group 1, experimental group 2, and experimental group 3 were compared, a significant difference in favor of experimental group 1 was observed compared to experimental group 3. When experimental group 2 was compared to experimental group 3, a significant difference was determined in favor of experimental group 2. It was determined that there is no significant difference between experimental group 1 and experimental group 2. However, when looking at the mean total scores in Figure 2, it was found that experimental group 1 scored higher than experimental group 2. The research results indicate that the phonological awaereness educaiton program is more effective than the digital game-based educaiton program. These findings are in line with relevant research in the leterature. The active participation of the children in the implementation process of the education program applied to experimental group 1, the flexibility of the education process, and the attempt to increase interaction with large and small group studies may have led to the emergence of this difference.

Shifflet et al. (2020) examined the effects of educational materials (traditional materials) and tablet-based applications on preschool children's phonological awareness skills. Traditional materials were used in the education given to one experimental group, while tablet-based applications were applied in the education given to the other experimental group. They found that the phonological awareness skills of children in both experimental groups improved. They also found that there was no significant difference between the experimental groups. In the study conducted by Shifflet et al. (2020), materials such as cubes, magnetic letters, and unifix were used. In the study, children's books, finger plays, rhymes, songs, natural objects, toys, and activities were prepared in a gamified way in the implementation of the phonological awareness educational program. The difference between the Shifflet et al. (2020) study and the study can be attributed to the materials used. Goffredo et al. (2016) developed a platform called 'En Plein' to support preschool children's phonological awareness skills, which includes activities related to rhyme awareness, initial/final syllable recognition, and syllable blending skills. The study consisted of an experimental group that used the platform and a control group that received regular phonological awareness education. The research found that the phonological awareness education program was more effective than the digital game-based education program, which can be considered a more conventional program than the digital game-based education program. There are differences in the research results. Goffredo et al. (2016) noted that boys and girls participating in their research were highly motivated. The ability of children to create their avatars and move simultaneously with the avatar may have increased their motivation, leading to more accurate answers. This could be a source of the differences in research results. Hillman and Marshall (2009) emphasized the importance of intrinsic

motivation for achieving desired results in skills targeted through digital use and for these results to be lasting.

According to the results, experimental group 1, which received the phonological awareness education program, scored higher in all ELST sub-dimensions than the control group. Experimental group 1 had a higher average score in rhyme word matching, initial sound recognition, and overall scores than experimental groups 3 and 2. Experimental group 2 was found to have higher scores in rhyme word matching, generating words starting with the prompting sound, and overall scores compared to experimental group 3. When the final test findings for all sub-dimensions of ELTKC and ELST are examined among the experimental groups, the phonological awareness education program applied to experimental group 1 is more effective.

In their study, Elimelech and Aram (2020) developed a digital syllable game for children aged 5-7 to support early literacy skills. The research was conducted with three different groups, including two experimental groups and a control group. According to the study results, the group provided with both visual and auditory support and the group provided with only auditory support performed better than the other groups in Word syllabication and phonological awareness areas. The group with visual and auditory support received higher letter knowledge, phonological awareness, and word spelling scores than the unsupported and control groups. Children in both experimental groups scored higher in letter knowledge than in the control group. In the syllabication of words, the groups receiving auditory and visual-auditory support related to word syllabication scored higher in the final test than those receiving only auditory support and the control group. When examining the research, there appears to be a similarity between the study and the implementation process. In the conducted research, there are three different experimental groups: a phonological awareness education program group, a digital game-based phonological awareness education program group, and a digital game-supported phonological awareness education program group. Looking at the stimuli provided to the experimental groups in the research, the group where the digital game-supported phonological awareness education program was implemented was exposed to more stimuli. In comparison, the other groups were exposed to fewer stimuli. In this regard, Elimelech and Aram (2020) concluded their research that the group with more stimuli (visual and auditory support) scored higher in all areas compared to the other groups. However, in the conducted research, it is observed that among the groups that received high scores, the group that underwent the phonological awareness education program performed better. Therefore, there is no similarity in the research results. However, in the study conducted by Elimelech and Aram (2020), when examining the group with auditory support and the group without support compared to the other groups, it was observed that the group with auditory support scored higher. These results are similar to the digital game-supported phonological awareness education program applied in experimental group 2 and the digital game-based phonological awareness education program applied in experimental group 3, which scored higher than the control group. Likewise, it is similar to the digital game-based phonological awareness education program received by experimental group 3, which scored higher than the control group. The digital game applied to experimental group 3 includes visual and auditory elements. Providing two different sensory supports during the education process may have resulted in similar research results. At the same time,

the fact that the education program applied to experimental group 2 included both digital games and activities using different method techniques, appealing to children's tactile, visual, and auditory senses and interacting directly with objects and peers may have caused similar results.

Comparison of the Experimental Groups with the Control Group

When comparing experimental group 1, which received the phonological awareness education program, with the control group based on the ELTKC and ELST final test results, a significant increase in the scores of experimental group 1 is observed. The results obtained are similar to the research results in the literature. The results obtained are similar to the research results in the literature. For example, In a study conducted by Bolduc (2009) examining the effect of a music education program on preschool children's phonological awareness skills, an experimental music education program was applied to the study group, while the control group followed the official curriculum of the ministry. The study results concluded that the experimental music education program was more effective in developing participant children's phonological awareness skills. On the other hand, in a study conducted by Bayraktar and Temel (2014), they found that the pre-literacy education program implemented with different activities and methods was influential on the reading comprehension, mechanical reading, and writing skills of children in the experimental group as they progressed to primary school education. Kelly et al. (2019), in their research examining the effect of a phonological awareness educational program with traditional materials on children aged 3.5 to 5.5 years, found that the word and alphabet awareness of the experimental group improved. The use of music as both an activity type and a method, starting from different activities and using materials children encounter in their daily lives, maybe the reason for the similarity with the research findings.

When comparing experimental group 2, which received the digital game-supported phonological awareness education program, with the control group based on the ELTKC and ELST final test results, it can be observed that experimental group 2 showed a better score increase than the control group. In a study conducted by Jadán-Guerrero et al. (2020), using both natural materials and a digital program with children aged 5-6 who had phonological awareness issues, it was concluded that the applied program had a positive impact on the development of children's phonological awareness skills. Weiss et al. (2022) prepared an online reading camp education program for five-year-old children. The research included five-year-old children from the experimental group, who participated in the online reading camp program, and the control group, who did not receive the education. The research results showed that children in the experimental group improved their phonological awareness and decoding skills compared to the control group. The study aligns with previous research findings, and it can be said that the implemented digital game-supported phonological awareness education program contributes to children's phonological awareness skills.

Again, when comparing experimental group 3, which received the digital game-based phonological awareness education program, with the control group based on the ELTKC and ELST final test results, it can be observed that experimental group 3 scored higher than the control group. The implemented digital game has contributed to the children's phonological awareness skills. While there may not have been specific

content related to syllable skills in the digital game used in the study by Li et al. (2020), the progress in syllable skills observed in children could be attributed to their age. However, the difference in the study could also be because the digital game content was prepared following the ranking of phonological awareness skills and the educational program was developed based on a spiral approach. In their research, Da Silva et al. (2022) examined the impact of digital games on literacy skills in groups of 4-year-old children consisting of four experimental groups and a control group. They found that the experimental groups scored higher than the control group in phonological, syllable, and rhyme awareness skills. Oliva-Maza et al. (2021) developed an adventure-themed digital game to support phonological awareness skills in four preschool-age children aged 5-6. During the game process, children were required to complete different tasks related to phonological awareness in each section. At the end of the process, it was concluded that the adventure-themed digital game supported children's phonological awareness skills. The research aligns with the findings of previous studies.

Examination of the Persistence of the Experimental Groups

When examining the ELTKC retention test results, it was concluded that the education programs applied to experimental group 1, 2, and 3 had a retention effect. Additionally, an increase in the scores of experimental group 2 was observed in the ELTKC retention test results. According to the ELST retention test findings, it was determined that the education programs applied to experimental group 2 and 3 had a retention effect. Furthermore, an increase in the ELST retention test results was observed in experimental group 3. According to the ELST retention test findings, it was concluded that there was a decrease in the scores of experimental group 1. However, it was also determined that there was no significant difference in the ELST subtests between the final test and the retention test for experimental group 1. Therefore, while the education program applied to experimental group 1 had a retention effect on the ELST subtests, it did not affect the overall results. The applied phonological awareness, digital game-supported phonological awareness, and digital game-based phonological awareness education programs had a retention effect. In their study, Çetin (2019) found that the early literacy program they implemented had a lasting effect on phonological awareness skills. Similarly, Dinler and Cevher Kalburan (2021) concluded that the poetry-supported education program had a lasting effect on word, rhyme, initial/final sound, and syllable awareness skills. Van der Kooy-Hofland et al. (2012) determined that a computer game focused on names and sounds had a lasting effect on children's phonological awareness skills. Sá et al. (2022) developed a digital game to support phonological awareness skills and found that the developed digital game had a lasting effect on children's phonological awareness skills. It is essential to review and repeat information and experiences periodically. Repeating with both new and previously learned words helps achieve lasting learning (Beers et al., 2010). The research findings and the existing literature in the field support the results of the conducted study.

Upon analyzing the data obtained from the research, it can be concluded that the phonological awareness education program (experimental group 1), the digital game-supported phonological awareness education program (experimental group 2), and the digital game-based phonological awareness education program (experimental group 3) are all effective programs. Additionally, the findings suggest that the most effective

among the implemented education programs is the phonological awareness education program (experimental group 1). Furthermore, it can be interpreted that the digital game-supported phonological awareness education program (experimental group 2) is more effective than the digital game-based phonological awareness education program (experimental group 3). In the education program applied to Experimental Group 1, the inclusion of materials from children's daily lives and the use of different materials, intensive interaction with peers and the educator, high participation in the education process with their thoughts and actions, and the use of different methods and techniques may have made this education program more effective. The lack of interaction with both the game and peers during the digital game process, the absence of an adventure element, and the fact that it was played on a smart board instead of a tablet may have caused experimental group 3 to be less effective than the other groups.

Implications

With the development of artificial intelligence and different digital applications, studies can be carried out by developing games for phonological awareness that can be played with individual tools, where children are more active, can make changes in the game, and interact more with their peers and the game. For digital games to support the education process, games that match the theme used in the education process and the game theme can be used. Family-based or family-participated interventions using digital games for phonological awareness development can also be investigated. The effect of phonological awareness skills on cognitive and social development can be investigated. Moreover, the research process for implementing various educational programs can explore children's participation, willingness, and motivation. During the education process, phonological awareness skills (combining sounds, making sounds, etc.) can be included, taking into account the developmental characteristics of children.

No financial support was received from any institution for this research. Technical support was obtained for developing the digital game created for the research. We want to express our gratitude to Sekizdesekiz Group for their contributions.

Statement of Responsibility

Both authors contributed to the work. The initial conceptualization and drafting of the original text were carried out by Göle. Methodological design was carried out by Göle and Temel and analysis by Göle. Data, Software and Visualization are carried out to the lake. Writing, Revision by Göle and Temel, and Editing by Temel and Göle.

Conflicts of Interest

There are no conflicts of interest among the authors in this research. Furthermore, there are no conflicts of interest with any institution.

Author Bios:

Mehmet Oğuz Göle completed his undergraduate education at İnönü University and his master's and doctoral education in the Department of Early Childhood Education at Gazi University. He currently works as a lecturer in the Department of Child Development at Afyon Vocational School. His areas of expertise include quality standards in early childhood education, phonological awareness in early childhood, and early childhood digital games.

Zeynep Fulya Temel completed her undergraduate and master's education in the Department of Child Development and Education at Hacettepe University, and she completed her doctoral education in the Department of Child Health and Education at the same university. She currently serves as a Professor at the Faculty of Education at Gazi University. Her areas of expertise include language development, family education and learning, early childhood education programs, and different approaches in early childhood education.

References

Amorim, A. N., Jeon, L., Abel, Y., Felisberto, E. F., Barbosa, L. N., & Dias, N. M. (2020). Using escribo play video games to improve phonological awareness, early reading, and writing in preschool. *Educational Researcher*, *49*(3), 188-197. https://doi.org/10.3102/0013189X20909824

Arias, M. P. (2023). *Phonological Awareness Games-A Take Home Kit*. Greensboro College.

Arnold, D. H., Chary, M., Gair, S. L., Helm, A. F., Herman, R., Kang, S., & Lokhandwala, S. (2021). A randomized controlled trial of an educational app to improve preschoolers' emergent literacy skills. *Journal of Children and Media*, *15*(4), 457-475. https://doi.org/10.1080/17482798.2020.1863239

Arnott, L. (2013). Are we allowed to blink? Young children's leadership and ownership while mediating interactions around technologies. *International Journal of Early Years Education*, *21*(1), 97-115. https://doi.org/10.1080/09669760.2013.772049

Bayraktar, V., & Temel, Z. F. (2014). Okuma yazmaya hazırlık eğitim programı'nın çocukların okuma yazma becerilerine etkisi [The effect of the program of readiness education on the skills on reading-writing skills]. *Hacettepe University Journal of Education*, 29(3), 08-22.

Becker, K. (2017). *Choosing and using digital games in the classroom* (pp. 175-214). Springer.

Beers, C. S., Beers, J. W., & Smith, J. O. (2010). *A principal's guide to literacy teaching.* Guilford Press.

Bennett, M., & Parise, M. (2014). Computer technology: Possibilities and pitfalls. *Childhood Education*, *90*(5), 365-369. https://doi.org/10.1080/00094056.2014.952579

Bentin, S., Hammer, R., & Cahan, S. (1991). The effects of aging and first grade schooling on the development of phonological awareness. *Psychological Science*, *2*(4), 271-275. https://doi.org/10.1111/j.1467-9280.1991.tb00148.x

Bolduc, J. (2009). Effects of a music programme on kindergartners' phonological awareness skills 1. *International Journal of Music Education*, *27*(1), 37-47. https://doi.org/10.1177/0255761408099063

Bolduc, J., & Lefebvre, P. (2012). Using nursery rhymes to foster phonological and musical processing skills in kindergarteners. *Creative Education*, *3*(4), 495-502. https://doi.org/10.4236/ce.2012.34075

Busink, R. (1997). Reading and phonological awareness: What we have learned and how we can use it. *Literacy Research and Instruction*, *36*(3), 199-215. https://doi.org/10.1080/19388079709558239

Büyüköztürk, Ş. (2013). *Sosyal bilimler için veri analizi el kitabı*. Pegem Akademi.

Büyüköztürk, Ş., Kılıç-Çakmak, E., Akgün, Ö. E., Karadeniz, Ş., & Demirel, F. (2014). *Bilimsel araştırma yöntemleri* (Geliştirilmiş 16. Baskı). Pegem Akademi.

Cameron, M. (2023). *The effects of a computerized intervention on kindergarten students' reading skills* [Unpublished doctoral dissertation]. Mount Saint Vincent University.

Cardoso-Martins, C., Mesquita, T. C. L., & Ehri, L. (2011). Letter names and phonological awareness help children to learn letter–sound relations. *Journal of Experimental Child Psychology*, *109*(1), 25-38. https://doi.org/10.1016/j.jecp.2010.12.006

Casalis, S., & Colé, P. (2009). On the relationship between morphological and phonological awareness: Effects of training in kindergarten and in first-grade reading. *First Language*, *29*(1), 113-142. https://doi.org/10.1177/0142723708097484

Castles, A., & Coltheart, M. (2004). Is there a causal link from phonological awareness to success in learning to read? *Cognition*, *91*(1), 77-111. https://doi.org/10.1016/S0010-0277(03)00164-1

Çetin, A. (2019). *Erken okuryazarlık eğitimi eğitimi erken okuryazarlık ve matematik hataları üzerine bilgisayar incelemesi* [Investigation of the effect of early literacy skills educational program on early literacy and early mathematics skills. [Unpublished doctoral dissertation]. Hacettepe University.

Chambers, B., Cheung, A. C., & Slavin, R. E. (2016). Literacy and language outcomes of comprehensive and developmental-constructivist approaches to early childhood education: A systematic review. *Educational Research Review*, *18*, 88-111. https://doi.org/10.1016/j.edurev.2016.03.003

Check, J., & Schutt, R. K. (2012). *Research methods in education*. Sage Publications.

Christie, J. F., Enz, B., & Vukelich, C. (2011). *Teaching language and literacy: Preschool through the elementary grades*. Longman.

Cohen, L., Manion, L., & Morrison, K. (2017). *Research methods in education*. Routledge.

Da Silva, G. C., Rodrigues, R. L., Amorim, A. N., Mello, R. F., & Neto, J. R. O. (2022). Game learning analytics can unpack Escribo play effects in preschool early reading and writing. *Computers and Education Open*, *3*, 100066. https://doi.org/10.1016/j.caeo.2021.100066

Degé, F., & Schwarzer, G. (2011). The effect of a music program on phonological awareness in preschoolers. *Frontiers in Psychology*, *2*, 124. https://doi.org/10.3389/fpsyg.2011.00124

De Jong, P. F. (2007). Phonological awareness and the use of phonological similarity in letter–sound learning. *Journal of Experimental Child Psychology*, *98*(3), 131-152. https://doi.org/10.1016/j.jecp.2007.06.003

Del Egido, M. G. V. (2023). *The effects of pitch-focused and rhythm-focused music programs on preschool phonological awareness, alphabet knowledge, and concept of word awareness* [Unpublished doctoral dissertation]. The George Washington University.

Dinler, H., & Cevher Kalburan, N. (2021). The effect of poetry focused supportive educational program to preschool children's phonological awareness. *Kastamonu Education Journal, 29*(5), 1034-1051. https://doi.org/10.24106/kefdergi.715624

Elimelech, A., & Aram, D. (2020). Using a digital spelling game for promoting alphabetic knowledge of preschoolers: The contribution of auditory and visual supports. *Reading Research Quarterly*, *55*(2), 235-250. https://doi.org/10.1002/rrq.264

Elmonayer, R. A. (2013). Promoting phonological awareness skills of Egyptian kindergarteners through dialogic reading. *Early Child Development and Care*, *183*(9), 1229-1241. https://doi.org/10.1080/03004430.2012.703183

Ferrari, L., & Addessi, A. R. (2014). A new way to play music together: The Continuator in the classroom. *International Journal of Music Education*, *32*(2), 171-184. https://doi.org/10.1177/0255761413504706

Franc, B., & Subotic, V. (2015). Differences in phonological awareness of five-year-olds from Montessori and regular program preschool institutions. In *U: Researching paradigms of childhood and education conference book of selected papers (2nd Symposium: Child language and culture)* (pp. 12-20).

Goffredo, M., Bernabucci, I., Lucarelli, C., Conforto, S., Schmid, M., Nera, M. M., Lopez, L., D'Alessio, T., & Grasselli, B. (2016). Evaluation of a motion-based platform for practicing phonological awareness of preschool children. *Journal of Educational Computing Research*, *54*(5), 595-618. https://doi.org/10.1177/0735633115626881

Gökmen, G. S. (2007). 4; 0-6; 0 yaş (48-72 aylar) arasındaki çocukların ad-eylem kullanımları [Noun verb usage of 4;0-6;0 (the 48-72 month children) year old children]. *Dil Dergisi*, *137*, 18-33.

Görgün, E. (2020). *Resimli çocuk kitaplarının felsefi kavramlar bakımından incelenmesi* [Unpublished master thesis]. Gazi University.

Goswami, U. (2003). Early phonological development and the acquisition of literacy. Dickinson, D. K., & Neuman, S. B. (Eds.). *Handbook of Early Literacy Research*. Guilford Press.

Harper, L. J. (2011). Nursery Rhyme Knowledge and Phonological Awareness in Preschool Children. *Journal of Language and Literacy Education*, *7*(1), 65-78.

Haynie, W. J. (1997). Effects of anticipation of test on delayed retention learning. *Journal of Technology Education*, 9 (1), 20-30.

Hempenstall, K. (2003). Phonemic awareness: What does it mean? *Education Oasis*, *201*.14.08.2021https://pact.tarleton.edu/TCERT/Content/Documents/Phonemic%20Awareness.pdf

Hillman, M., & Marshall, J. (2009). Evaluation of digital media for emergent literacy. *Computers in the Schools*, *26*(4), 256-270. https://doi.org/10.1080/07380560903360186

Hogan, T. P., Catts, H. W., & Little, T. (2005). The relationship between phonological awareness and reading: Implications for the assessment of phonological awareness. *Language, Speech, and Hearing Services in the Schools, 36,* 285–293. https://doi.org/10.1044/0161-1461(2005/029)

Incognito, O., & Pinto, G. (2023). Longitudinal effects of family and school context on the development on emergent literacy skills in preschoolers. *Current Psychology*, *42*(12), 9819-9829. https://doi.org/10.1007/s12144-021-02274-6

Jadán-Guerrero, J., Ramos-Galarza, C., los Angeles Carpio-Brenes, M. D., Calle-Jimenez, T., Salvador-Ullauri, L., & Nunes, I. L. (2020). Phonological awareness intervention and basic literacy skill development with Kiteracy-PiFo. In *International Conference on Applied Human Factors and Ergonomics*, 319-325. Springer, Cham. https://doi.org/10.1007/978-3-030-51369-6_43

Justice, L. M., & Vukelich, C. (Eds.). (2008). *Achieving excellence in preschool literacy instruction*. Guilford Press.

Karaman, G., & Aytar, A. G. (2016). Erken okuryazarlık becerilerini değerlendirme aracı'nın (EOBDA) geliştirilmesi [Development of early literacy skills assessment tool (ELST)]. *Mersin University Journal of the Faculty of Education*, *12*(2), 516-541. https://doi.org/10.17860/efd.02080

Kargın, T., Güldenoğlu, B., & Ergül, C. (2017). Anasınıfı çocuklarının erken okuryazarlık beceri profili: Ankara örneklemi [Early literacy profile of kindergarten children: Ankara sample]. *Journal of Special Education*, *18*(1), 61-87. https://doi.org/10.21565/ozelegitimdergisi.299868

Kartal, G., & Terziyan, T. (2016). Development and evaluation of game-like phonological awareness software for kindergarteners: JerenAli. *Journal of Educational Computing Research*, *53*(4), 519-539. https://doi.org/10.1177/0735633115608397

Keklik, S. (2010). Türkçede 0-6 yaş çocuklarına öğretilmesi gereken en sık kullanılan 1200 kelime. *Türkiye Sosyal Araştırmalar Dergisi*, *3*, 1-28.

Keklik, S. (2011). Türkçede on bir yaşına kadar çocuklara öğretilmesi gereken, birleşim gücü yüksek ilk bin kelime [The first thousand words in turkish which can easily be combined should be taught until eleven year old children]. *ODÜ Sosyal Bilimler Araştırmaları Dergisi (ODÜSOBİAD)*, *2*(4), 80-95.

Kelly, C., Leitão, S., Smith-Lock, K., & Heritage, B. (2019). The effectiveness of a classroom-based phonological awareness program for 4–5-year-olds. *International Journal of Speech-Language Pathology*, *21*(1), 101-113. https://doi.org/10.1080/17549507.2017.1400589

Kostelnik, M. J., Soderman, A. K., Whiren, A. P., & Rupiper, M. L. (2019). *evelopmentally appropriate curriculum: Best practices in early childhood education*. Pearson Education.

Kozminsky, L., & Kozminsky, E. (1995). The effects of early phonological awareness training on reading success. *Learning and Instruction*, *5*(3), 187-201. https://doi.org/10.1016/0959-4752(95)00004-M

Lefebvre, P., Trudeau, N., & Sutton, A. (2011). Enhancing vocabulary, print awareness and phonological awareness through shared storybook reading with low-income preschoolers. *Journal of Early Childhood Literacy*, *11*(4), 453-479. https://doi.org/10.1177/1468798411416581

Lenhart, J., Suggate, S. P., & Lenhard, W. (2022). Shared-reading onset and emergent literacy development. *Early Education and Development*, *33*(4), 589-607. https://doi.org/10.1080/10409289.2021.1915651

Lennon, M., Pila, S., Flynn, R., & Wartella, E. A. (2022). Young children's social and independent behavior during play with a coding app: Digital game features matter in a 1: 1 child to tablet setting. *Computers & Education*, *190*, 104608. https://doi.org/10.1016/j.compedu.2022.104608

Lennox, S. (2014). The potential of poetry for early literacy learning: Why, how and what? *Practically Primary*, *19*(3), 21-24.

Li, Y., Chen, X., Li, H., Sheng, X., Chen, L., Richardson, U., & Lyytinen, H. (2020). A computer-based Pinyin intervention for disadvantaged children in China: Effects on Pinyin skills, phonological awareness, and character reading. *Dyslexia*, *26*(4), 377-393. https://doi.org/10.1002/dys.1654

Lim, C. T., & Chew, F. P. (2018). Using poems to increase phonological awareness among children. In *Issues and Trends in Interdisciplinary Behavior and Social Science: Proceedings of the 6th International Congress on Interdisciplinary Behavior and Social Sciences (ICIBSoS 2017), July 22-23, 2017, Bali, Indonesia* (p. 33). CRC Press.

Luna, S. M. (2021). Challenging norms in pre-kindergarten curriculum by listening to young children: Pre-service teachers' lessons in phonological awareness. *International Critical Childhood Policy Studies Journal*, *8*(2), 80-91.

Martins, M. A., & Silva, C. (2001). Letter names, phonological awareness and the phonetization of writing. *European Journal of Psychology of Education*, *16*, 605-617.

McClarty, K. L., Orr, A., Frey, P. M., Dolan, R. P., Vassileva, V., & McVay, A. (2012). A literature review of gaming in education. *Gaming in education*, *1*(1), 1-35.

Memoğlu-Süleymanoğlu, H. (2014a). *Türkçenin sıklık sözlüğü. Hatiboğlu Yayınları*.

Memoğlu-Süleymanoğlu, H. (2014b). *Türkçenin ters sıklık sözlüğü. Hatiboğlu Yayınları*.

Mertala, P., & Meriläinen, M. (2019). The best game in the world: Exploring young children's digital game–related meaning-making via design activity. *Global Studies of Childhood*, *9*(4), 275-289. https://doi.org/10.1177/2043610619867701

Mertens, D. M. (2019). *Research and evaluation in education and psychology: Integrating diversity with quantitative, qualitative, and mixed methods*. Sage Publications.

Mihai, A., Friesen, A., Butera, G., Horn, E., Lieber, J., & Palmer, S. (2015). Teaching phonological awareness to all children through storybook reading. *Young Exceptional Children*, *18*(4), 3-18. https://doi.org/10.1177/1096250614535221

Ministry of National Education [MoNE]. (2013). 36-72 aylık çocuk için okul öncesi eğitim programı. 20.05.2021 https://tegm.meb.gov.tr/dosya/okuloncesi/ooproram.pdf

Moritz, C., Yampolsky, S., Papadelis, G., Thomson, J., & Wolf, M. (2013). Links between early rhythm skills, musical training, and phonological awareness. *Reading and Writing, 26,* 739-769. https://doi.org/10.1007/s11145-012-9389-0

Morrison, T. G., & Wilcox, B. G. (2012). *Developing literacy: Reading and writing to, with, and by children*. Pearson Higher Ed.

Moyer, P. S., Bolyard, J. J., & Spikell, M. A. (2002). What are virtual manipulatives? *Teaching children mathematics*, *8*(6), 372-377. https://doi.org/10.5951/TCM.8.6.0372

National Reading Panel. (2000). *Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction: Reports of the subgroups*. National Institute of Child Health and Human Development, National Institutes of Health.

Newby, P. (2010). *Research methods for education*. Pearson Education.

Nichols, W. D., Rasinski, T. V., Rupley, W. H., Kellogg, R. A., & Paige, D. D. (2018). Why poetry for reading instruction? Because it works! *The Reading Teacher*, *72*(3), 389-397. https://doi.org/10.1002/trtr.1734

Oliva-Maza, A., Ayuso-Escuer, N., Coma-Roselló, T., & Torres-Moreno, E. F. (2021). Mystery of the runaway letrabytes: Inclusive assessment of phonological awareness with tangible gamification. *IEEE Revista Iberoamericana de Tecnologias del Aprendizaje*, *16*(4), 424-432.

Özen, Y., & Gül, A. (2007). Sosyal ve eğitim bilimleri araştırmalarında evren-örneklem sorunu [Population-sampling issue on social and educational research studies]. *Journal of Kazım Karabekir Education Faculty*, *15*, 394-422.

Patria, L. (2023, June). Improving students 'motivation in learning phonics through games. In *Proceedings of the 2nd International Conference on Advances in Humanities, Education and Language, ICEL 2022*, 07–08 November 2022, Malang, Indonesia.

Phillips, B. M., Clancy-Menchetti, J., & Lonigan, C. J. (2008). Successful phonological awareness instruction with preschool children: Lessons from the classroom. *Topics In Early Childhood Special Education*, *28*(1), 3-17. https://doi.org/10.1177/0271121407313813

Pila, S., Aladé, F., Sheehan, K. J., Lauricella, A. R., & Wartella, E. A. (2019). Learning to code via tablet applications: An evaluation of Daisy the Dinosaur and Kodable as learning tools for young children. *Computers & Education*, *128*, 52-62. https://doi.org/10.1016/j.compedu.2018.09.006

Pogiatzi, M., Bardoutsou, I., Lavidas, K., & Komis, V. (2022). Interviewing preschool children in Greece about their usage of mobile devices at home. *SN Social Sciences*, *2*(10), 215. https://doi.org/10.1007/s43545-022-00522-5

Pullen, P. C., & Justice, L. M. (2003). Enhancing phonological awareness, print awareness, and oral language skills in preschool children. *Intervention in School and Clinic*, *39*(2), 87-98. https://doi.org/10.1177/10534512030390020401

Puolakanaho, A., Poikkeus, A. M., Ahonen, T., Tolvanen, A., & Lyytinen, H. (2003). Assessment of three-and-a-half-year-old children's emerging phonological awareness in a computer animation context. *Journal of Learning Disabilities, 36*(5), 416-423. https://doi.org/10.1177/00222194030360050301

Redig, J. (2018). Does daily exposure through the use of nursery rhymes better support students' phonological awareness skill development? [Unpublished Master Thesis]. Northwestern Collage, Iowa.

Rideout, V. (2011). Zero to eight: Children's media use in America. Common Sense Media.

Rowe, M. L., Kirby, A. L., Dahbi, M., & Luk, G. (2023). Promoting language and literacy skills through music in early childhood classrooms. *The Reading Teacher, 76*(4), 487-496. https://doi.org/10.1002/trtr.2155

Ryan, R. M., Rigby, C. S., & Przybylski, A. (2006). The motivational pull of video games: A self-determination theory approach. *Motivation and Emotion, 30*, 344-360. https://doi.org/10.1007/s11031-006-9051-8

Sá, M., Sa-Couto, P., & Lousada, M. (2022). Phonological awareness digital program: a randomized controlled study. *Revista de Investigación en Logopedia, 12*(1). https://doi.org/10.5209/rlog.77402

Şahin Kamışlı, G., Mengü, G., Bayramoğlu, İ., & Kemaloğlu, Y. (2015). Konuşma odyometrisi ve çocuklar için kelime listeleri geliştirilmesi üzerine bir derleme [A review on speech audiometry and developing pediatric word list]. *Türkiye Klinikleri, 8*(2), 13-25.

Şahin, L. (2019). *Okul eğitim dönemi erken dönem okur yazarlık öğretimine öğretmene yönelik bakışın incelenmesi* [Master's dissertation]. Yakın Doğu Üniversitesi.

Saracho, O., & Spodek, B. (2002). Influences on Early Childhood Curriculum Development. Saracho, O., & Spodek, B. (Eds.). (2002). *Contemporary perspectives on early childhood curriculum*. IAP.

Savaş, M., & Turan, M. (2011). Okulöncesi altı yaş grubu öğrencilerin sahip oldukları kelime servetinin incelenmesi [Assessment of word power of age group of six students in the pre-scohool time]. *Education Sciences, 6*(1), 841-859.

Selwyn, N. (2016). *Education and technology: Key issues and debates*. Bloomsbury Publishing.

Schuele, C. M., & Boudreau, D. (2008). Phonological awareness intervention: Beyond the basics. *Language, Speech, And Hearing Services in Schools*, *39*(1), 3-20. https://doi.org/10.1044/0161-1461(2008/002)

Sheridan, S., Williams, P., Sandberg, A., & Vuorinen, T. (2011). Preschool teaching in Sweden–a profession in change. *Educational Research, 53*(4), 415-437. https://doi.org/10.1080/00131881.2011.625153

Shifflet, R., Mattoon, C., & Bates, A. (2020). Using tablets in a prekindergarten classroom to foster phonological awareness. *International Research in Early Childhood Education*, *10*(1), 1-20.

Shuler, C. (2012). *iLearn II: An analysis of the education category on apple's app store*. Joan Ganz Cooney Center.

Stadler, M. A., & McEvoy, M. A. (2003). The effect of text genre on parent use of joint book reading strategies to promote phonological awareness. *Early Childhood Research Quarterly*, *18*(4), 502-512. https://doi.org/10.1016/j.ecresq.2003.09.008

Stephen, C., & Plowman, L. (2014). Digital lay. In Brooker L., Blaise M., & Edwards S. (Eds.), *SAGE handbook of play and learning in early childhood*. SAGE.

Sucena, A., Silva, A. F., & Marques, C. (2023). Reading skills promotion: Results on the impact of a preschool intervention. *Frontiers in Education, 7*, 1076630.

Symmonds, J. M. (2020). *Impact of quality shared reading experiences on the phonological awareness skills of preschoolers* [Unpublished doctoral dissertation]. The William Paterson University.

Topbaş, S. (1996). Sesbilgisi açısından dil edimin süreci. *Dilbilim Araştırmaları Dergisi*, *7*, 295-309.

Topbaş, S. (2006). Türkçe sesletim-sesbilgisi testi: Geçerlik-güvenirlik ve standardizasyon çalışması, *Türk Psikoloji Dergisi, 21*(58), 39-56.

Torgesen, J. K., & Mathes, P. G. (1998). *What every teacher should know about phonological awareness*. Florida Department of Education, Division of Schools and Community Education, Bureau of Instructional Support and Community Services.

Van der Kooy-Hofland, V. A., Bus, A. G., & Roskos, K. (2012). Effects of a brief but intensive remedial computer intervention in a sub-sample of kindergartners with early literacy delays. *Reading and Writing*, *25*, 1479-1497. https://doi.org/10.1007/s11145-011-9328-5

Van Goch, M. M., Verhoeven, L., & McQueen, J. M. (2017). Trainability in lexical specificity mediates between short-term memory and both vocabulary and rhyme awareness. *Learning and Individual Differences*, *57*, 163-169. https://doi.org/10.1016/j.lindif.2017.05.008

Vinter, A., Bard, P., Lukowski-Duplessy, H., & Poulin-Charronnat, B. (2022). A comparison of the impact of digital games eliciting explicit and implicit learning processes in preschoolers. *International Journal of Child-Computer Interaction*, *34*, 100534. https://doi.org/10.1016/j.ijcci.2022.100534

Weiss, Y., Yeatman, J. D., Ender, S., Gijbels, L., Loop, H., Mizrahi, J. C., Woo, B.Y., & Kuhl, P. K. (2022). Can an online reading camp teach 5-year-old children to read? *Frontiers in Human Neuroscience*, 52. https://doi.org/10.3389/fnhum.2022.793213

Xiong, Z., Liu, Q., & Huang, X. (2022). The influence of digital educational games on preschool Children's creative thinking. *Computers & Education*, *189*, 104578. https://doi.org/10.1016/j.compedu.2022.104578

Yopp, H. K., & Yopp, R. H. (2000). Supporting phonemic awareness development in the classroom. *The Reading Teacher*, *54*(2), 130-143.

Zevenbergen, R. (2007). Digital natives come to preschool: Implications for early childhood practice. *Contemporary Issues in Early Childhood*, *8*(1), 19-29. https://doi.org/10.2304/ciec.2007.8.1.19

Journal of Theoretical Educational Science, 17(1), 236-256, January 2024
Kuramsal Eğitimbilim Dergisi, 17(1), 236-256, Ocak 2024
[Online]: http://dergipark.org.tr/akukeg
DOI number: http://doi.org/10.30831/akukeg.1316813

Development of Recycling Attitude Scale

Geri Dönüşüm Tutum Ölçeğinin Geliştirilmesi

Mustafa DERMAN* **Şeyda GÜL**** **Mehmet ERKOL*****

Received: 19 June 2023 **Research Article** **Accepted:** 18 January 2024

ABSTRACT: The aim of this research is to develop a scale for determining pre-service teachers' attitudes towards recycling. The research was carried out using a quantitative research design. The research was carried out in the education faculties of two state universities located in the east and west of Türkiye. 284 pre-service teachers (62 males, 222 females) who were selected by the easily accessible sampling method participated in the research. During the scale development process, an item pool consisting of 56 statements was created. The draft scale, which was prepared after the expert opinion, was applied to the pre-service teachers. After validity and reliability analysis, three factors with 32 items were obtained, which explained 42.456% of the total variance. The factors were named "Responsibility and Behavior," "Consciousness and Awareness," and "Economic Value", respectively. The Cronbach"s Alpha coefficient was calculated as 0.893 for the "Responsibility and Behavior" factor, 0.785 for the "Consciousness and Awareness" factor, and 0.801 for the "Economic Value" factor. In addition, Cronbach's Alpha was calculated as 0.884 for the overall scale. According to the findings, it can be said that this scale developed in the research is sufficient in terms of validity and reliability. However, in future studies, similar scales with confirmatory factor analysis can be developed and applied to large samples to examine participants' attitudes toward recycling.

Keywords: Recycling, attitude, scale development, pre-service teacher.

ÖZ: Bu araştırmanın amacı öğretmen adaylarının geri dönüşüm konusundaki tutumlarını belirlemeye yönelik bir ölçek geliştirmektir. Araştırma nicel araştırma deseni ile yürütülmüştür. Araştırma, Türkiye'nin doğusunda ve batısında yer alan iki devlet üniversitesinin eğitim fakültelerinde gerçekleştirilmiştir. Araştırmaya kolay ulaşılabilir örnekleme yöntemiyle seçilen 284 (62 erkek, 222 kadın) öğretmen adayı katılmıştır. Ölçek geliştirme sürecinde 56 ifadeden oluşan bir madde havuzu oluşturulmuştur. Uzman örüşü sonrasında hazırlanan taslak ölçek öğretmen adaylarına uygulanmıştır. Geçerlik ve güvenirlik analizleri sonucunda toplam varyansın %42.456'sını açıklayan 32 maddelik 3 faktör elde edilmiştir. Faktörler sırasıyla "Sorumluluk ve Davranış", "Bilinç ve Farkındalık" ve "Ekonomik Değer" olarak adlandırılmıştır. Cronbach's Alpha katsayısı "Sorumluluk ve Davranış" faktörü için .893, "Bilinç ve Farkındalık" faktörü için .785 ve "Ekonomik Değer" faktörü için .801 olarak hesaplanmıştır. Ayrıca ölçeğin geneli için Cronbach's Alpha değeri .884 olarak hesaplanmıştır. Ulaşılan bulgulara göre, araştırmada geliştirilen bu ölçeğin geçerlik ve güvenirlik açısından yeterli olduğu söylenebilir. Bununla beraber gelecek araştırmalarda DFA'nın da yapıldığı benzer ölçekler geliştirilebilir ve ölçek geniş örneklemlere uygulanarak katılımcıların geri dönüşüme yönelik tutumları incelenebilir.

Anahtar kelimeler: Geri dönüşüm, tutum, ölçek geliştirme, öğretmen adayı**.**

* *Corresponding Author*: Dr, Atatürk University, Erzurum, Türkiye, mderman@atauni.edu.tr, https://orcid.org/0000-0002-5263-4262
** Prof. Dr., Atatürk University, Erzurum, Türkiye, seydagul@atauni.edu.tr, https://orcid.org/0000-0003-4005-2158
*** Assoc. Prof., Afyon Kocatepe University, Afyonkarahisar, Türkiye, merkol@aku.edu.tr, https://orcid.org/0000-0002-1363-1894

Citation Information
Derman, M., Gül, Ş., & Erkol, M. (2024). Development of recycling attitude scale. *Kuramsal Eğitimbilim Dergisi [Journal of Theoretical Educational Science], 17*(1), 236-256.

ISSN: 1308-1659

The world population is increasing rapidly. It is estimated that if the rate of population growth continues like this, the world population will reach over 10 billion by the end of the 21st century. Due to the rapidly increasing world population, the requirements for the resources people use are increasing, and people meet the majority of these needs from natural resources. However, natural resources in the world are limited. Therefore, to consume fewer natural resources, the raw materials meet human requirements by recycling (UNEP, 2016). However, recycling is not at the desired level worldwide. Waste and recycling are among the problems the world is facing today. All countries have been making great efforts to reduce the amount of waste and to encourage people to recycle. Various services such as picking up waste from home (door-to-door collection), paying fees to encourage recycling and recycling bins have been implemented for people to acquire habits about recycling. The main purpose of these implementations is to encourage people to recycle (Darby & Obara, 2005; Haj-Salem & Al-Hawari, 2021; Tonglet et al., 2004). However, the tendency of people around the world to recycle remains very low (Haj-Salem & Al-Hawari, 2021). Hence, to enable people to participate in the recycling process actively and to improve recycling behavior, many countries, especially European countries, are aiming to improve their waste management systems and recycle more products (Haj-Salem & Al-Hawari, 2021; Zaikova et al., 2022).

Recycling does not only include the collection of waste and making them useful. Natural resources are also protected through recycling. Because the raw material source is provided by natural resources, in other words, raw materials that cannot be met through recycling are supplied directly from natural sources. Without recycling, natural resources will be further depleted. For example, metal ores are consumed more due to the lack of recycling of metals. Similarly, if plastic products are not recycled, more hydrocarbons and fossil fuels will be used to produce plastic products (Lamma, 2021; Maddox, 1972). Therefore, Recycling provides benefits such as energy saving and the protection of nature for people. For example, recycling paper instead of cutting trees will save energy to cut many trees and protect green nature (Chan & Bishop, 2013). However, today, despite the intense efforts related to recycling, people also produce a large amount of waste. Even in developed countries, this rate is very high (Darby & Obara, 2005; Elgaaied, 2012; Zaikova et al., 2022). Because waste, instead of being recycled, is deposited in dumpsites in cities. Dumpsites are common in many countries around the world, including developed countries. Most of the garbage is collected in dumpsites without being recycled. It has been revealed that there is a high connection between the population density of people and dumpsites. The majority of the population is located 10 kilometers from the dumpsites. There are also a lot of dumpsites in areas where the population lives densely (UNEP, 2016). According to OECD data, the waste generation of countries and the recycling of wastes differ on the basis of countries. However, it has been stated that the consumption and waste production rate has increased depending on the country's economic growth. However, in low-economic countries, the rate of waste collected in open dumps without processing is around 93% (OECD G20, 2021). Comparisons between Türkiye and OECD countries have determined that the collection of waste in landfills is above the average of OECD countries in Türkiye. However, it has a value below that of OECD countries in terms of recycling (OECD, 2021).

Sustainability is based on the consumption of natural resources to reach future generations. All the resources in the world are limited and people have difficulties finding resources due to over-consumption. However, nature is excessively destroyed in order to meet the resource requirement. The natural balance is destroyed and nature cannot renew itself. Therefore, countries are developing policies in line with the understanding of sustainability such as protecting resources in a balanced way, reducing waste, increasing recycling, and increasing people's awareness about reducing waste, reusing, and recycling (Phulwani et al., 2020). Furthermore, in order to reduce the effect on and damage to the environment, eco-friendly products have been started to be produced in recent years. Eco-friendly products are products that do not harm the environment or cause minimal damage when they are thrown into the environment (Siddique & Hossain, 2018). These products, which can be biodegradable easily and in a short time in nature, help to protect the environment and nature (Tseng et al., 2018). In this context, there is a trend towards a green economy to protect nature and cause less damage. Green economy is based on reducing carbon emissions, using natural resources effectively, expanding renewable energy sources, recycling, and reducing harm to nature (UNEP, 2023). Therefore, by transitioning to a green economy, governments establish all policies on environmentally friendly and sustainable foundations (Ferrão et al., 2014). With a green economy, socio-economic and environmental improvements are addressed as a whole. In other words, socio-economic developments are addressed holistically with environmental risks (Kumar, 2017; Pahle et al., 2016). Less use of natural resources and increased investments in recycling allow ecosystems to renew themselves. In addition, the green economy allows the development of policies to reduce environmental risks (Kumar, 2017). Because of the destruction of natural resources, economic crises occur throughout the world. Natural resources are the main sources of economies. Recycling reduces the need to allocate financial resources to the purchase and use of raw materials. Recycling not only conserves natural resources but also meets the need for raw materials and provides jobs. (EEA, 2011; Ferrão et al., 2014). However, unless people's environmental awareness and awareness levels are increased, the efforts of governments will not be enough. Therefore, society needs to be aware of the protection of the environment and the sustainable consumption of natural resources. For this reason, researchers have focused on researching the factors that affect individuals' sustainable consumer behavior over 40 years (Antonetti & Maklan, 2014).

Theoretical Background on Recycling Behavior

Researchers have developed various theories to determine the factors that affect people's behavior on consumption, waste, and recycling (Chan & Bishop, 2013). One of the most important theories explaining this situation today is the *theory of planned behavior* (TPB). According to this theory, people's behavior choices are based on *attitude, subjective norms*, and *perceived behavioral control* (Wang et al., 2019). In recent years, the TPB model has been widely preferred to determine the factors that affect recycling and waste-related behaviors. Because it is one of the best models to explain the factors affecting behavior (Wang et al., 2019), another model *norm activation model* (NAM) is proposed by Schwartz (1973). The essence of this model is based on personal norms (Schwartz, 1973). Some researchers try to identify the factors

that affect recycling behavior by integrating different models (integration of NAM–TPB model) (Onwezen et al., 2013). Each model revealed by scholars can help people develop environmentally friendly behaviors (Stern, 2000).

Many factors can affect the choice of behavior. The factors that have a high effect on the individual's environmentally friendly behavior intention constitute the most important factor directing the behavior of individuals. In other words, it is the level of intention to perform the behavior (Ajzen, 1991). The variables in recycling behavior may differ according to the research area. These include beliefs, behaviors, economics, motivation, sense of responsibility, education level, age, gender, and so on. Scholars have been trying to reveal how these variables affect each other (Wang et al., 2019). Because many disciplines (e.g., education, psychology, and economy) are related to recycling, conduct research. Each field is trying to determine the factors that affect recycling behaviors from its own point of view (Hornik et al., 1995). Furthermore, some researchers have tried to explain the link between identity and recycling behavior. The relationship between self-identity and behavioral intention is revealed using the TPB model (Fekadu & Kraft, 2001; Whitmarsh & O'Neill, 2010). When the causes of the factors affecting environmental behaviors are revealed, environmental education will enable people to have environmentally friendly behaviors because the most important aim of environmental education is to improve people's awareness of the use of natural resources in a sustainable way (Torkar & Bogner, 2019). Numerous studies have been conducted to determine attitudes and behaviors toward the environment (Chawla & Cushing, 2007; Cohen, 1993; Elgaaied, 2012; Hsu & Lin, 2015; Passafaro & Livi, 2017; Quoquab et al., 2020; Robinson & Read, 2005; Szczytko et al., 2019; Vilkaite-Vaitone & Jeseviciute-Ufartiene, 2021). However, one of the factors not emphasized by the researchers is the effect of economic concern on environmental behavior. Because the relationship between economic concerns and environmental behavior has not been fully examined, this situation has been mostly addressed in the context of three values (egoistic, altruistic, and biospheric) (Snelgar, 2006; Swami et al., 2010). However, there were not enough scale items related to economic concerns. For example, the scale used by Whitmarsh and O'Neill (2010), items-buying eco-friendly products, saving behavior at home, using renewable energy sources at home, using public transportation, and participating in environmental activities- were included to determine the relationship between identity and pro-environmental behaviors. But only one item- "*Drive economically"*- is related to the economic concern. In the Tonglet et al. (2004) study, the environmental behaviors of individuals regarding recycling were investigated in terms of the *theory of planned behavior* (TPB). The section about situational factors includes one item in terms of economic behavior. Recycling is expressed as a waste of money. Another item- "*Recycling saves money"*- is included in the factor of consequences of recycling. In a similar study, economic concern was measured with the item "*recycling saves energy"* (Davies et al., 2002). To determine the effect of emotional variables on recycling behaviors, items on economic concerns were not included, whereas related to environmental concerns, results of recycling, and municipalities were used (Elgaaied, 2012). In some items used within the personnel norms (e.g., we should protect the environment), recycling action may take place due to both economic and other reasons (Onwezen et al., 2013). A similar situation (e.g., ecological problems are the consequences of my actions) was also used in the study by

Antonetti and Maklan (2014). In recent study by Zaikova et al. (2022), the effects of economic incentives on recycling behavior have begun to be included in the *theory of planned behavior* (TPB) as a variable. As a result of the research, in this case different countries (Russia and Finland) were compared, it was determined that economic incentives did not have an effect on "*waste source-separation intention"*. But Wang et al. (2019) pointed out that economic motivation has an effect on recycling behavior.

Many studies have been conducted using similar scales. Various items related to economic concerns are as follows;

Table 1

Items Used in terms of Economic Concerns

Recycling saves energy	(Davies et al., 2002; Schoeman & Rampedi, 2022; Tonglet et al., 2004; Vining & Ebreo, 1992)
Recycling saves money	(Schoeman & Rampedi, 2022; Tonglet et al., 2004)
Drive economically	(Whitmarsh & O'Neill, 2010)
Recycling programmes are a waste of money	(Tonglet et al., 2004)
Monetary benefit in waste sorting	(Zaikova et al., 2022)
Reduced waste collection fees when I sort the waste	(Zaikova et al., 2022)
Refund of the cost of packaging (for example, plastic bottles) if I recycle it	(Zaikova et al., 2022)
Wash and reuse dishcloths rather than buying them new	(Barr, 2007)
Waste separation can help the country to reduce pollution control costs, we should do it	(Xu et al., 2017)
Humans have the right to modify the natural environment to suit their needs	(Vining & Ebreo, 1992)
Plants and animals exist primarily to be used by humans	(Vining & Ebreo, 1992)

As reviewed in the literature given above, it has been determined that the economic concerns in the recycling behavior scales are not under-represented. Although there were no economic concerns in the scales used by the researchers, some items were included in relation to the economic concerns (Davies et al., 2002; Tonglet et al., 2004; Whitmarsh & O'Neill, 2010; Xu et al., 2017; Zaikova et al., 2022). However, Wang et al. (2019) and Zaikova et al. (2022) examined the effect of the economic variables on recycling behavior is examined in terms of economic motivation and incentives. It is important to develop scales that include economic variables. Because people are destroying nature in order to gain more profit, many reasons, such as over-consumption of natural resources, destruction of biodiversity, and conversion of forest areas into agricultural lands, are the behaviors that people do in order to live more economically. However, it is possible to eliminate these negative behaviors, first of all, with education. Environmental education enables students to take an active role in environmental issues in the formal education process. In fact, environmental education aims to increase the

knowledge of students and to develop responsible behavior (Chawla & Cushing, 2007). In the Belgrade Charter (1975), the objectives of environmental education are grouped under 6 headings (*awareness, knowledge, attitude, skills, evaluation ability*, and *participation*). It is emphasized as the objectives aimed to be gained by students in the education process (UNESCO-UNEP, 1975). At this point, teachers, who are the most important part of education, have great duties. In this respect, it is another important point of this research to develop a scale by referring to the opinions of prospective teachers who will be teachers in the future. Therefore, the aim of this research is to develop a scale for determining pre-service teachers' attitudes towards recycling.

Method

This section includes the research method, research group, ethical processes, scale development process, and measures for validity and reliability.

Research Design

Since this research aims to develop a scale, it can be said that it is based on a quantitative research design. The concept of scale essentially shows the mathematical qualities of measurement results (Gül & Sözbilir, 2015). It is also used in many fields of behavioural sciences, such as education and psychology, to collect information in terms of the targeted person or persons, system, subject, or content (Yurdugül, 2005).

Research Group

The investigation was conducted in the education faculties of two state universities located in the east and west of Türkiye. A total of 284 prospective teachers (62 males, 222 females) participated in the research, and convenience sampling was used as a sampling method. In addition, since both factor analysis and structural equation modeling are analysis types that require large samples, as well as to increase the power of the analytical model tested, the sample size was tried to be large (MacCallum et al., 1999). Thus, according to Bryman and Cramer (2001), the number of samples was determined according to the rule that "the sample must be at least five times the number of items in the scale." In order to perform the study, the data was collected in the 2022-2023 academic year. In the data collection process, the participation of teacher candidates on a voluntary basis was ensured. In addition, the personal information of the participants was not asked when filling out the scales, and the applications were carried out based on confidentiality. The demographic characteristics of the participants included in the study are presented in Table 2.

Table 2

Demographic Characteristics of the Participants

Participants	Gender	1st grade	2nd grade	3rd grade	4th grade	Total
University 1	Male	6 (17.1%)	4 (21.1%)	6 (12.8%)	10 (27.8%)	26 (19.0%)
	Female	29 (82.9%)	15 (78.9%)	41 (87.2%)	26 (72.2%)	111 (81.0%)
University 2	Male	-	16 (20.0%)	6 (17.1%)	14 (46.7%)	36 (24.5%)
	Female	2 (1.8%)	64 (57.7%)	29 (26.1%)	16 (14.4%)	111 (75.5%)

As seen in Table 2, the study includes sample groups from two different universities. Of the participants in the first group, 111 (81.0%) were female and 26 (19.0%) were male. Of the participants in the second group, 111 (75.5%) were female and 36 (24.5%) were male.

Ethical Procedures

All procedures were confirmed by researchers in accordance with the ethical standards of Ataturk University (date: 21.03.2023, reference no: E-56785782-050.02.04-2300106373).

The Development Process of the Scale

The principles suggested by Karakoç and Dönmez (2014) were taken into consideration during the scale development process. According to Karakoç and Dönmez (2014), if a new scale is to be developed, the following steps should be followed: (1) Conducting a literature review on the subject. (2) To determine the format for the measurement method and create an item pool accordingly. (3) To seek expert comment for the item pool, thus assessing content and face validity. (4) Make a trial application. (5) After the trial application, validity and reliability analysis of the scale and creating the final form of the scale. The development process of the scale is presented below.

On the other hand, according to DeVellis and Thorpe (2021), the items to be prepared at this stage should reflect the concept (phenomenon) being researched. For this reason, each item must comply with the structure of the implicit variable. In other words, the items written should not go beyond the conceptual framework established in the first step. It can create an item pool with inductive (asking open-ended questions to the target audience) and deductive methods (literature review) (Evci & Aylar, 2017). Since a limited sample was reached in this research and voluntary participation was ensured, an item pool was created only by the deductive method.

For the scale, the relevant literature was previously examined and an item pool containing 58 statements was created. The statements in the item pool were examined in terms of language, comprehensibility and content by four academicians, one of whom is an expert in science education and three of whom are experts in biology education. After the opinions of experts, two expressions that were not understandable or thought to be similar were discarded from the item pool. Thus, the number of items was reduced to 56, and then it was decided to perform 5-point Likert-type scoring in line with the evaluation criteria of the scale items in the literature (Shure & Spivack, 1982). The 5-point Likert scale consists of these points: (1) Strongly Disagree; (2) Disagree; (3) Partially Agree; (4) Agree; (5) Strongly Agree. Thus, the draft scale was made ready for implementation.

Validity and Reliability Measures

The scales whose reliability and validity have been tested and found to be sufficient will provide valid data for the person applying the scale (Ercan & Kan, 2004). For this reason, the validity and reliability of the scale developed in the study were tested.

In determining the validity of the measurement tool, content validity, criterion validity, construct validity and face validity are checked (Ercan & Kan, 2004). On the

other hand, the content validity of a scale can be examined in two ways: logically and statistically (Ercan & Kan, 2004). In this study, the content validity and face validity of the scale were ensured logically by examining the literature and consulting four academicians, one of whom is an expert in science education and three of whom are experts in biology education. Criterion validity examines the future or current relationship between the scores obtained from the scale and the determined criterion in order to determine the effectiveness of the scale. Since there is no similar scale in the literature that includes all the dimensions of this scale, criterion validity was not examined in this research. Factor analysis was performed for the construct validity of the test. Although different methods such as norm-referenced test, test-retest method, parallel-forms method and methods of interval consistency were used in the reliability studies of the scale (Ercan & Kan, 2004), in this research, the reliability of the scale was tested by calculating the Cronbach Alpha Reliability Coefficient, one of the internal consistency methods.

Data were collected simultaneously for each grade in the classroom under the supervision of a teacher. Before answering the measurement tool, students were informed about the purpose of the study and it was stated that their personal information would be kept confidential by the researchers.

In the study, it was also considered to ensure external validity in determining the number of samples. According to Büyüköztürk et al. (2013), one of the factors affecting external validity is the sampling effect. In other words, people selected from a limited area are unlikely to represent people elsewhere. In this case, the result applies to the individuals included in the research. Therefore, the study tried to increase external validity by including teacher candidates from two universities in the east and west of Türkiye.

Results

Preliminary Analysis and Validity Studies

To ensure the validity of the draft scale, the content, face, and construct validity were used. As stated before, the items in the scale for content and face validity were submitted to the opinions of four experts, and a draft scale of 56 items was prepared. Furthermore, to construct the validity of the scale, factor analysis was performed. Exploratory Factor Analysis (EFA) and Confirmatory Factor Analysis (CFA) techniques are widely used in scale adaptation or development processes for construct validity. If there is no known relationship between the scale items, EFA is recommended. On the contrary, if a tested relationship, specified factors and items grouped under them are detected, CFA is recommended (Bandalos & Finney, 2010; Büyüköztürk, 2002). In this study, since there was no theoretical information among the scale items, that is, it was not known exactly how many factors there were among the items and which items measured which factors, only EFA was performed. Additionally, the literature states that CFA should be carried out with a different sample (Fabrigar et al., 1999). CFA could not be performed due to the study's insufficient sample size. SPSS 20.0 was used for all analyses.

Before performing the analyses, the data were examined in terms of missing values and no missing data was found. On the other hand, kurtosis and skewness

coefficients were calculated in order to evaluate the normality assumptions. According to the analysis results, the kurtosis value was calculated as 1.07, while the skewness value was calculated as -.29. According to Leech et al. (2005), if the distribution is normal, the skewness and kurtosis coefficients should be between -1 and 1. It is stated that if the skewness coefficient is between -1 and 1, the kurtosis coefficient may be between -2 and 2, and if the kurtosis coefficient is between -1 and 1, the skewness coefficient may be between -2 and 2. Therefore, these findings showed that the data set met normal distribution assumptions.

Item Analysis

In the item analysis stage, the corrected item-total correlations were calculated. In the literature, these values are accepted as the measure of the effect of each item on the scale. Additionally, if the item-total correlation value is less than 0.25 or negative, it is recommended that it be removed from the scale (Gul, 2017). Accordingly, 10 items (I11, I13, I34, I40, I41, I43, I44, I50, I52, I55) was eliminated from the scale. After this stage, the number of remaining items was reduced to 46 (Table 3).

Table 3

The Corrected Item-Total Correlations of the Items in the Scale

Item	Item-total correlation	*t*	Item	Item-total correlation	*t*	Item	Item-total correlation	*t*
I1	0.355	-6.559	I20	0.480	-8.465	I39	0.420	-6.535
I2	0.282	-4.914	I21	0.355	-7.030	**I40**	**0.203**	-3.502
I3	0.308	-4.633	I22	0.508	-10.863	**I41**	**0.123**	-3.453
I4	0.436	-6.319	I23	0.551	-10.528	I42	0.502	-8.126
I5	0.309	-4.764	I24	0.333	-5.493	**I43**	**0.056**	-1.401*
I6	0.351	-4.529	I25	0.523	-8.655	**I44**	**-0.099**	1.270*
I7	0.320	-5.738	I26	0.293	-4.781	I45	0.315	-5.131
I8	0.464	-5.798	I27	0.496	-7.862	I46	0.501	-9.025
I9	0.461	-5.783	I28	0.327	-6.417	I47	0.293	-4.602
I10	0.419	-5.197	I29	0.563	-9.734	I48	0.271	-4.897
I11	**0.131**	-2.552	I30	0.366	-6.691	I49	0.503	-9.278
I12	0.518	-7.856	I31	0.467	-8.213	**I50**	**-0.336**	4.593
I13	**0.180**	-2.975	I32	0.445	-7.611	I51	0.534	-10.451
I14	0.432	-7.392	I33	0.338	-6.643	**I52**	**0.163**	-3.536
I15	0.523	-9.969	**I34**	**0.074**	-1.933*	I53	0.464	-7.516
I16	0.494	-9.744	I35	0.410	-6.910	I54	0.528	-8.030
I17	0.444	-7.581	I36	0.428	-9.283	**I55**	**0.125**	-2.574
I18	0.431	-7.354	I37	0.451	-9.685	I56	0.441	-6.131
I19	0.475	-8.358	I38	0.536	-9.868			

* $p>.05$

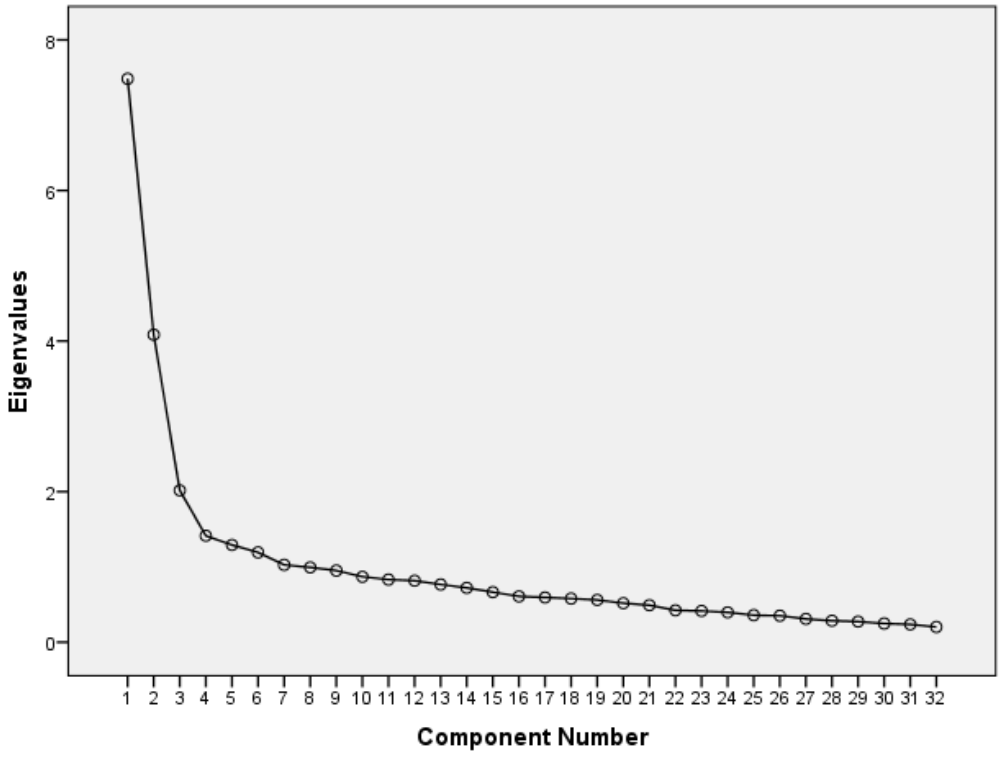
Factor Analysis

After item analysis, Exploratory Factor Analysis (EFA) was performed for the construct validity of the scale. In factor analysis, it is necessary to test the suitability of the data for factor analysis. In addition, Bartlett's test of sphericity, which is an indicator of the multivariate normal distribution of the data, is expected to be significant. As a result of the preliminary analysis, the KMO sample fit coefficient was .876 (>.60), and for Bartlett's test of sphericity was $\chi 2$=5669.095 (p<.05). These findings showed that the data was suitable for performing exploratory factor analysis.

Principal component analysis was used as a factorization technique in EFA. However, in determining the number of factors, and the factors with an eigenvalue greater than one were considered. The varimax rotation technique was used to interpret the factors. Factor loads are known as the correlation value between the item and the structure in the scale (Kocaman & Cumaoğlu, 2014). Accordingly, items with a factor load value below 0.30 were removed from the scale in analysis. Another criterion considered in item elimination is the removal of items that fall under more than one factor and differ less than 0.10 from the scale (Kocaman & Cumaoğlu, 2014). Accordingly, nine items (I2, I17, I25, I28, I29, I35, I37, I39 and I54) were removed from the scale in the first factor analysis. After ongoing analysis, three factors whose eigenvalues were greater than one were obtained, which explained 42.456% of the total variance (Table 4). Additionally, a scree plot was also made (Figure 1). It refers to a factor between two points on the scree plot graph (Büyüköztürk, 2002).

Figure 1

Scree Plot Graphic



When Figure 1 is examined, the 3-factor structure is clearly seen.

Table 4

The Results of EFA

		Rotated component matrix		
Item no	Items	Factor 1	Factor 2	Factor 3
I20	I pay attention to buying eco-friendly products.	0.778		
I19	I collect the garbage I see while walking.	0.753		
I21	I pay attention to whether there is an eco-friendly statement on the packaging.	0.734		
I23	I encourage people to recycle.	0.716		
I24	I actively participate in local or global activities related to recycling and waste.	0.697		
I16	I separate waste according to recycling types.	0.653		
I33	I collect the garbage I see and throw it in the trash to set an example for those around me.	0.644		
I32	I do not take the trash that other people throw on the ground and throw it in the trash.	0.637		
I22	News about recycling on TV or radio intrigues me.	0.637		
I15	I throw waste into recycling bins.	0.619		
I18	I warn those who throw garbage on the ground.	0.568		
I14	I know which waste goes into which color waste bin.	0.542		
I30	I try to throw the waste I see around me in the trash.	0.482		
I8	Throwing cigarette butts on the street does not pose an environmental problem.		0.746	
I10	Liquid oils become harmless after use in food, so there is no need for recycling.		0.696	
I9	There is no harm in pouring the waste oil into the sink.		0.694	
I12	Recycling is the responsibility of the municipality. Therefore, we do not need to make an effort to recycle waste.		0.659	
I6	Wastes thrown into water resources such as rivers, lakes and seas do not pose a danger to humans.		0.602	
I7	Reusing a product has no environmental benefit.		0.520	
I3	Recycling helps protect nature.		0.475	
I1	I am aware of the importance of recycling.		0.458	
I4	Biodiversity can be protected by recycling.		0.432	
I5	There is no relationship between sustainable development and recycling.		0.387	
I49	Recycling should only be done for expensive and valuable products.			0.760
I51	People can use natural resources to generate more income.			0.754

I46	It makes more sense to buy a new product instead of the expenses incurred in the recycling process.			0.718
I42	Recycling has no economic benefit. Therefore, there is no need for recycling.			0.626
I45	Recycling process wastes money.			0.583
I48	Due to the cost of the waste in the recycling process, it should be disposed of away from settlements.			0.557
I53	Recycling helps conserve resources.			0.515
I47	Recycling is a costly process.			0.459
I26	I throw away an old product and buy a new one.			0.435
Eigen value (Total=13.585%)		5.817	3.920	3.848
Explained variance (Total=42.456%)		18.179	12.251	12.026

As shown in Table 4, the scale, which includes 32 items and under three factors, were named "Responsibility and Behavior (Factor 1)", "Consciousness and Awareness (Factor 2)" and "Economic Value (Factor 3)", respectively. These items explain about 42.5% of the total variance. Furthermore, the load values of Factor 1, which includes 13 items, are between 0.778 and 0.482. The item loadings of Factor 2, which includes 10 items, are between 0.746 and 0.387. The item loadings of the last factor, which is called Factor 3, are between 0.760 and 0.435.

In the research, correlations between the subscales formed after the factor analysis were examined (Table 5).

Table 5

The Pearson Correlations among Subscales

	Factor 1	Factor 2	Factor 3
Factor 1	1	0.324**	0.210**
Factor 2	0.324**	1	0.468**
Factor 3	0.210**	0.324**	0.324**

**p<.001

Table 5 shows that the relationships between the subscales are all positively related.

Reliability Analysis

In the research, the reliability of the overall scale and its three factors were determined by calculating Cronbach's Alpha coefficient. According to the findings, this value (α) was 0.893 for the "Responsibility and Behavior" factor, 0.785 for the "Consciousness and Awareness" factor and 0.801 for the "Economic Value" factor. Moreover, Cronbach's Alpha was calculated as 0.884 for the overall scale. As Özdamar (2004) stated, coefficients 0.60 and greater indicate *good reliability* and high consistency among scale items. In this context, scale and its factors are quite reliable. The final form of the scale is in Appendix 1.

Discussion and Conclusion

Recycling practices are increasing significantly due to the necessity of solving environmental problems that have reached extreme levels today. For this reason, research on related factors affecting recycling has gained momentum in recent years (Momoh & Oladebeye, 2010; Šorytė & Pakalniškienė, 2021; Tam et al., 2018). When the literature on recycling is examined, it shows that although attitudes toward recycling play an important role in recycling and environmental education, valid and reliable measurement tools are limited in determining students' attitudes (Ugulu, 2015). In this research, an attitude scale was developed to determine the attitudes of pre-service teachers toward recycling. The scale development process was basically performed in four stages (item pool, item analyses, EFA, and reliability analyses). Firstly, an item pool, which includes 56 items, was created based on examining the literature and according to expert opinions. Item analyses were made on the collected data, and then a three-factor, "Responsibility and Behavior," "Consciousness and Awareness," and "Economic Value" was obtained.

In the literature, in the field of social sciences, the variance explained according to EFA results is expected to be between 0.40 and 0.60 for structures with more than one factor (Çokluk et al., 2012; Türkan & Çeliköz, 2018). In the research, the explained variance of the scale was determined as 42.456%. Therefore, according to the EFA results, it was determined that the variance ratio explained by the items collected under three factors provided the desired variance ratio.

According to the EFA results, the difference between the load values of the items on different factors is expected to be over 0.10 (Can, 2013). According to the research findings, it was understood that the factor loadings of the items in the scale varied between 0.387 and 0.778 and that the items in the final version of the scale were not loaded on different factors. It is emphasized that the item load values in the scale should be above 0.30 (Çokluk et al., 2012) or above 0.32 (Tabachnick & Fidell, 2007). On the other hand, it can be said that the model, which consists of 32 items and three factors, was suitable both theoretically and statistically because these results obtained with EFA provide evidence that the scale has construct validity. The Cronbach Alpha internal consistency coefficients calculated within the scope of reliability studies showed that the scale was reliable. Accordingly, 13 items were collected in Factor 1 (Responsibility and Behavior), nine items in Factor 2 (Consciousness and Awareness), and Factor 3 (Economic Value). Therefore, the minimum score that can be obtained from the whole scale is 32, whereas the maximum score is 160.

When the findings obtained in the study are evaluated, it can be said that this scale, which is developed to determine the attitudes of teacher candidates toward recycling, has appropriate qualifications. Considering the cognitive, affective and behavioral dimensions of attitudes, it is thought that this structure, which is revealed by the developed scale, has features belonging to each sub-dimension of attitude and will help researchers in measuring these features. In addition, when the scales developed for recycling in the literature were examined, factors similar to or different from the scale in this research were determined. For example, in the "belief" factor determined in the scale development study conducted by Karatekin (2013), items were collected in a manner similar to the "consciousness and awareness" factor in our study. Again, the "interest and sensitivity" and "initiative and participation" factors are similar to the

"responsibility and behavior" factors in our study. Moreover, in the scale development study conducted by Avan et al. (2011), items were collected in the "environment-information" factor, similar to the "consciousness and awareness" factor in our study. Again, the "environment-emotion" and "environment-behavior" factors are similar to the "responsibility and behavior" factors in our study. In another study conducted by Taştepe (2017), an attitude scale was developed to determine the cognitive, affective, and behavioral attitudes of high school students regarding the reduction, reuse and recycling of packaging waste. As a result of the analysis, it was determined that the 10-item final scale had a two-factor structure, namely "Giving Emotional Reactions" and "Exhibiting Awareness and Appropriate Behavior," and explained 57.955% of the total variance. It can be said that these dimensions are partially similar to the "Responsibility and Behavior" and "Consciousness and Awareness" dimensions in our research.

On the other hand, unlike these studies, the "economic value" factor developed in our research can be considered as a new dimension that contributes to the literature. In addition, although the development studies of this scale were made with pre-service teachers, it is appropriate to measure the attitudes of university students from all branches.

Implications

The scale developed in the research has some limitations as well as its contributions to the literature. At this point, it seems appropriate to make the following recommendations for future research:

- By conducting similar scale development studies with different sample groups, different dimensions can be determined from the literature and this research.

- Although this research was conducted with teacher candidates in the education faculties of two universities, the number of samples remained a little low due to the online data collection process and the collection of data through voluntary participation. Hence, confirmatory factor analysis (CFA) was not performed. In factor analysis studies, CFA is performed to test the fit of the model after EFA. However, CFA needs to be performed with the data collected from a different sample. Due to the insufficient sample size in this study, CFA analysis could not be performed. Therefore, it is recommended that CFA analysis be performed in future research.

- Due to the lack of sample, the attitudes of teacher candidates towards recycling could not be measured. Therefore, in future studies, this scale can be applied to prospective teachers to examine their attitudes toward recycling and compare them in terms of different variables.

- Since the subject of recycling is very popular with the developing technology today, it always has an important news value for the media. In addition, the way the media organs handle the issue is extremely important, and sometimes, it can bring more harm than good (Uzbay, 2009). For this reason, it is suggested that the "effect of media" factor should be added to the scales that determine attitudes towards recycling.

As a result, it is thought that this scale will be an effective data collection tool in determining the attitudes of teacher candidates towards recycling. In this context, the use of the scale by researchers is thought to be an effective tool in revealing the views of samples with different views on recycling, which is of great importance not only in our country but also all over the world, and its use is recommended.

Acknowledgements

All authors would like to thank all students for their contribution to the process of administering the test.

Statement of Responsibility

All authors contributed to the study. Initial conceptualization, drafting of the original manuscript, and methodology were carried out by the first and second authors. Data collection was carried out by all authors.

Conflicts of Interest

The authors have no competing interests to declare that are relevant to the content of this article.

Author Bios:

Dr. Mustafa Derman (PhD) is a Dr. in Biology Education, Department of Secondary Science and Mathematics Education. The mean area of experience is in biology education and environmental education.

Dr. Şeyda Gül is a Associate Prof. in Department of Secondary Science and Mathematics Education. The mean area of experience is in biology education.

Dr. Mehmet Erkol is a Associate Prof. in the Department of Science Education, Afyonkocatepe University, Turkey. He has experience on environmental education and science learning.

References

Ajzen, I. (1991). The theory of planned behavior. *Organizational Behavior and Human Decision Processes, 50*, 179-211. https://doi.org/10.47985/dcidj.475

Antonetti, P., & Maklan, S. (2014). Feelings that make a difference: How guilt and pride convince consumers of the effectiveness of sustainable consumption choices. *Journal of Business Ethics, 124*(1), 117-134. https://doi.org/10.1007/s10551-013-1841-9

Avan, Ç., Aydınlı, B., Bakar, F., & Alboga, Y. (2011). Preparing attitude scale to define students‟ attitudes about environment, recycling, plastic and plastic waste. *International Electronic Journal of Environmental Education, 1*(3), 179-191.

Bandalos, D. L., & Finney, S. J. (2010). Factor analysis: Exploratory and confirmatory. In G. R. Hancock & R. O. Mueller (Eds.), *The reviewer's guide to quantitative methods in the social sciences* (pp. 93-114). Routledge.

Barr, S. (2007). Factors influencing environmental attitudes and behaviors: A U.K. case study of household waste management. *Environment and Behavior, 39*(4), 435-473. https://doi.org/10.1177/0013916505283421

Bryman, A., & Cramer, D. (2001). *Quantitative data analysis with SPSS release on for windows*. Routledge.

Büyüköztürk, Ş. (2002). Faktör analizi: Temel kavramlar ve ölçek geliştirmede kullanımı. *Kuram ve Uygulamada Eğitim Yöntemleri, 32*, 470-483.

Büyüköztürk, Ş., Akgün, Ö. E., Karadeniz, Ş., Demirel, F., & Kılıç, E. (2013). *Bilimsel araştırma yöntemleri*. Pegem Akademi.

Can, A. (2013). *SPSS ile bilimsel araştırma sürecinde nicel veri analizi*. Pegem Akademi.

Chan, L., & Bishop, B. (2013). A moral basis for recycling: Extending the theory of planned behaviour. *Journal of Environmental Psychology, 36,* 96-102. https://doi.org/10.1016/j.jenvp.2013.07.010

Chawla, L., & Cushing, D. F. (2007). Education for strategic environmental behavior. *Environmental Education Research, 13*(4), 437-452. https://doi.org/10.1080/13504620701581539

Cohen, S. (1993). For parents particularly: Reclaiming our earth: Recycling and conservation. *Childhood Education, 70*(1), 44-46. https://doi.org/10.1080/00094056.1993.10520983

Çokluk, Ö., Şekercioğlu, G., & Büyüköztürk, Ş. (2012). *Sosyal bilimler için çok değişkenli istatistik: SPSS ve LISREL uygulamaları*. Pegem Akademi.

Darby, L., & Obara, L. (2005). Household recycling behaviour and attitudes towards the disposal of small electrical and electronic equipment. *Resources, Conservation and Recycling, 44*(1), 17-35. https://doi.org/10.1016/j.resconrec.2004.09.002

Davies, J., Foxall, G. R., & Pallister, J. (2002). Marketing theory: An integrated model of recycling. *Marketing Theory, 2*(1), 29-113.

DeVellis, R. F., & Thorpe, C. T. (2021). *Scale development: Theory and applications (Applied social research methods)* (Fifth ed.), SAGE Publications.

EEA. (2011). *Earnings, jobs and innovation: The role of recycling in a green economy.* European Environment Agency Report.

Elgaaied, L. (2012). Exploring the role of anticipated guilt on pro-environmental behavior - a suggested typology of residents in France based on their recycling patterns. *Journal of Consumer Marketing, 29*(5), 369-377. https://doi.org/10.1108/07363761211247488

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

Evci, N., & Aylar, F. (2017). Ölçek geliştirme çalışmalarında doğrulayıcı faktör analizinin kullanımı. *Sosyal Bilimler Dergisi, 4*(10), 389-412. https://doi.org/10.16990/SOBIDER.3386

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.

Fekadu, Z., & Kraft, P. (2001). Self-identity in planned behavior perspective: Past behavior and its moderating effects on self-identity-intention relations. *Social Behavior and Personality, 29*(7), 671-686. https://doi.org/10.2224/sbp.2001.29.7.671

Ferrão, P., Ribeiro, P., Rodrigues, J., Marques, A., Preto, M., Amaral, M., Domingos, T., Lopes, A., & Costa, E. I. (2014). Environmental, economic and social costs and benefits of a packaging waste management system: A Portuguese case study. *Resources, Conservation and Recycling, 85,* 67-78. https://doi.org/10.1016/j.resconrec.2013.10.020

Gul, S. (2017). Development of an attitude scale to measure the undergraduate students' attitudes towards nanobiotechnology. *Journal of Science Education & Technology, 26*, 519-533.

Gül, Ş., & Sözbilir, M. (2015). Thematic content analysis of scale development studies published in the field of science and mathematics education. *Education and Science, 40*(178), 85-102.

Haj-Salem, N., & Al-Hawari, M. D. A. (2021). Predictors of recycling behavior: the role of self-conscious emotions. *Journal of Social Marketing, 11*(3), 204-223. https://doi.org/10.1108/JSOCM-06-2020-0110

Hornik, J., Cherian, J., Madansky, M., & Narayana, C. (1995). Determinants of recycling behavior: A synthesis of research results. *Journal of Socio-Economics, 24*(1), 105-127. https://doi.org/10.1016/1053-5357(95)90032-2

Hsu, J. L., & Lin, T. Y. (2015). Carbon reduction knowledge and environmental consciousness in Taiwan. *Management of Environmental Quality: An International Journal, 26*(1), 37-52. https://doi.org/10.1108/MEQ-08-2013-0094

Jöreskog, K., & Sörbom, D. (1993). *LISREL 8: Structural equation modeling with the simplis command language*. Scientific Software International, Inc.

Karakoç, F. Y., & Dönmez, L. (2014). Ölçek geliştirme çalışmalarında temel ilkeler. *Tıp Eğitimi Dünyası, 40*, 39-49.

Karatekin, K. (2013). Öğretmen adayları için katı atık ve geri dönüşüme yönelik tutum ölçeğinin geliştirilmesi: Geçerlik ve güvenirlik çalışması. *Uluslararası Avrasya Sosyal Bilimler Dergisi, 4*(10), 71-90.

Kocaman, O., & Cumaoğlu, G. K. (2014). Developing a scale for vocabulary learning strategies in foreign languages. *Education and Science, 39*(176), 293-303.

Kumar, P. (2017). Innovative tools and new metrics for inclusive green economy. *Current Opinion in Environmental Sustainability, 24*, 47-51. https://doi.org/10.1016/j.cosust.2017.01.012

Lamma, O. A. (2021). The impact of recycling on the environment. *Internat Ional Journal of Applied Research, 7*(11), 297-302. https://doi.org/10.1016/0361-3658(85)90009-8

Leech, N. L., Barrett, K. C., & Morgan, G. A. (2005). *SPSS for intermediate statistics: Use and interpretation*. (Second ed.). Taylor & Francis.

MacCallum, R. C., Widaman, K. F., Zang, S., & Hong, S. (1999). Sample size in factor analysis. *Psychological Methods, 4*, 84-99.

Maddox, J. (1972). Raw materials and the price mechanism. *Nature, 236*, 331-334.

Mainieri, T., Barnett, E. G., Valdero, T. R., Unipan, J. B., & Oskamp, S. (1997). Green buying: The influence of environmental concern on consumer behavior [Compra ecológica: la influencia de la preocupación por el medio ambiente en el comportamiento del consumidor]. *Journal of Social Psychology, 137*(2), 189-204.

Momoh, J. J., & Oladebeye, D. H. (2010). Assessment of awareness, attitude and willingnessof people to participate in household solid waste recycling programme in Ado-Ekiti, Nigeria. *Journal of Applied Sciences in Environmental Sanitation, 5*(1), 93-105.

OECD G20. (2021). Towards a more resource-efficient and circular economy. *OECD Publishing, Paris*, 1-53.

OECD. (2021). *OECD Economics Surveys: Turkey. January*, 14.

Onwezen, M. C., Antonides, G., & Bartels, J. (2013). The norm activation model: An exploration of the functions of anticipated pride and guilt in pro-environmental behaviour. *Journal of Economic Psychology, 39*, 141-153. https://doi.org/10.1016/j.joep.2013.07.005

Özdamar, K. (2004). *Paket programlar ile istatistiksel veri analizi 1 [Statistical data analysis with package programs]* (5. Ed.). Kaan Publishing.

Pahle, M., Pachauri, S., & Steinbacher, K. (2016). Can the Green Economy deliver it all? Experiences of renewable energy policies with socio-economic objectives. *Applied Energy, 179*, 1331-1341. https://doi.org/10.1016/j.apenergy.2016.06.073

Passafaro, P., & Livi, S. (2017). Comparing determinants of perceived and actual recycling skills: The role of motivational, behavioral and dispositional factors. *Journal of Environmental Education, 48*(5), 347-356. https://doi.org/10.1080/00958964.2017.1320961

Phulwani, P. R., Kumar, D., & Goyal, P. (2020). A systematic literature review and bibliometric analysis of recycling behavior. *Journal of Global Marketing, 33*(5), 354-376. https://doi.org/10.1080/08911762.2020.1765444

Popescu, S., Rusu, D., Dragomir, M., Popescu, D., & Nedelcu, Ș. (2020). Competitive development tools in identifying efficient educational interventions for improving pro-environmental and recycling behavior. *International Journal of Environmental Research and Public Health, 17*(1), 156. https://doi.org/10.3390/ijerph17010156

Quoquab, F., Mohammad, J., & Shahrin, R. (2020). Pro-environmental behavior in nutricosmetics product purchase context: Scale development and validation. *International Journal of Pharmaceutical and Healthcare Marketing, 14*(2), 217-250. https://doi.org/10.1108/IJPHM-04-2019-0033

Robinson, G. M., & Read, A. D. (2005). Recycling behaviour in a London Borough: Results from large-scale household surveys. *Resources, Conservation and Recycling, 45*(1), 70-83. https://doi.org/10.1016/j.resconrec.2005.02.002

Schoeman, D. C., & Rampedi, I. T. (2022). Drivers of household recycling behavior in the city of Johannesburg, South Africa. *International Journal of Environmental Research and Public Health, 19*(10), 6229. https://doi.org/10.3390/ijerph19106229

Schwartz, S. H. (1973). Normative explanations of helping behavior: A critique, proposal, and empirical test. *Journal of Experimental Social Psychology, 9*, 349-364.

Shure, M. B., & Spivack, G. (1982). Interpersonal problem-solving in young children: A cognitive approach to prevention. *American Journal of Community Psychology, 10*(3), 341-355.

Siddique, M. Z. R., & Hossain, A. (2018). Sources of consumers awareness toward green products and its impact on purchasing decision in Bangladesh. *Journal of Sustainable Development, 11*(3), 9-22. https://doi.org/10.5539/jsd.v11n3p9

Snelgar, R. S. (2006). Egoistic, altruistic, and biospheric environmental concerns: Measurement and structure. *Journal of Environmental Psychology, 26*(2), 87-99. https://doi.org/10.1016/j.jenvp.2006.06.003

Šorytė, D., & Pakalniškienė, V. (2021). Environmental attitudes and recycling behaviour in primary school age: the role of the school and parents. *Psichologija, 63*, 101-117.

Stern, P. C. (2000). Toward a coherent theory of environmentally significant behavior. *Journal of Social Issues, 56*(3), 407-424.

Swami, V., Chamorro-Premuzic, T., Snelgar, R., & Furnham, A. (2010). Egoistic, altruistic, and biospheric environmental concerns: A path analytic investigation of their determinants. *Scandinavian Journal of Psychology, 51*(2), 139-145. https://doi.org/10.1111/j.1467-9450.2009.00760.x

Szczytko, R., Stevenson, K., Peterson, M. N., Nietfeld, J., & Strnad, R. L. (2019). Development and validation of the environmental literacy instrument for adolescents. *Environmental Education Research, 25*(2), 193-210. https://doi.org/10.1080/13504622.2018.1487035

Tabachnick, B. G., & Fidell, L. S. (2007). *Using multivariate statistics*. Allyn & Bacon/Pearson Education.

Tam, V. W., Le, K. N., Wang, J. Y., & Illankoon, I. C. S. (2018). Practitioners recycling attitude and behaviour in the Australian construction industry. *Sustainability, 10*(4), 1212.

Taştepe, T. (2017). A study to develop an attitude scale for recycling among high school student. *Eğitim Kuram ve Uygulama Araştırmaları Dergisi, 3*(2), 01-13.

Tonglet, M., Phillips, P. S., & Read, A. D. (2004). Using the theory of planned behaviour to investigate the determinants of recycling behaviour: A case study from Brixworth, UK. *Resources, Conservation and Recycling, 41*(3), 191-214. https://doi.org/10.1016/j.resconrec.2003.11.001

Torkar, G., & Bogner, F. X. (2019). Environmental values and environmental concern. *Environmental Education Research, 25*(10), 1570-1581. https://doi.org/10.1080/13504622.2019.1649367

Tseng, M. L., Wong, W. P., & Soh, K. L. (2018). An overview of the substance of resource, conservation and recycling. *Resources, Conservation and Recycling, 136*, 367-375. https://doi.org/10.1016/j.resconrec.2018.05.010

Türkan, A., & Çeliköz, N. (2018). Ortaöğretim öğrencilerine yönelik özel ders eğilim ölçeğinin geçerlik ve güvenirlik çalışması. *Uluslararası Bilimsel Araştırmalar Dergisi, 3*(2), 398-410.

Ugulu, I. (2015). Development and validation of an instrument for assessing attitudes of high school students about recycling. *Environmental Education Research, 21*(6), 916-942.

UNEP. (2016). Global waste management outlook. In D. C. Wilson (Ed.), *Global waste management outlook*. https://doi.org/10.18356/765baec0-en

UNEP. (2023). *Green economy*. https://www.unep.org/regions/asia-and-pacific/regional-initiatives/supporting-resource-efficiency/green-economy

UNESCO-UNEP. (1975). *International workshop on environmental education; The Belgrade Charter: a framework for environmental education; 1975.* https://www.eusteps.eu/wp-content/uploads/2020/12/Belgrade-Charter.pdf 8.11.2022

Uzbay, İ. T. (2009). Ülkemizdeki temel sorunlar ve madde bağımlılığı ile mücadele. *Meslek İçi Sürekli Eğitim Dergisi, 21/22*, 73-79.

Vilkaite-Vaitone, N., & Jeseviciute-Ufartiene, L. (2021). Predicting textile recycling through the lens of the theory of planned behaviour. *Sustainability (Switzerland), 13*(20). https://doi.org/10.3390/su132011559

Vining, J., & Ebreo, A. (1992). Predicting recycling behavior from global and specific environmental attitudes and changes in recycling opportunities. *Journal of Applied Social Psychology, 22*(20), 1580-1607. https://doi.org/10.1111/j.1559-1816.1992.tb01758.x

Wang, B., Ren, C., Dong, X., Zhang, B., & Wang, Z. (2019). Determinants shaping willingness towards online recycling behaviour: An empirical study of household e-waste recycling in China. *Resources, Conservation and Recycling, 143*(July 2018), 218-225. https://doi.org/10.1016/j.resconrec.2019.01.005

Whitmarsh, L., & O'Neill, S. (2010). Green identity, green living? The role of pro-environmental self-identity in determining consistency across diverse pro-environmental behaviours. *Journal of Environmental Psychology, 30*(3), 305-314. https://doi.org/10.1016/j.jenvp.2010.01.003

Xu, L., Ling, M., Lu, Y., & Shen, M. (2017). Understanding household waste separation behaviour: Testing the roles of moral, past experience, and perceived policy effectiveness within the theory of planned behaviour. *Sustainability (Switzerland), 9*(625), 1-27. https://doi.org/10.3390/su9040625

Yurdugül, H. (2005). *Ölçek geliştirme çalışmalarında kapsam geçerliği için kapsam geçerlik indekslerinin kullanılması*. XIV. Ulusal Eğitim Bilimleri Kongresi, Pamukkale Üniversitesi Eğitim Fakültesi, Denizli.

Zaikova, A., Deviatkin, I., Havukainen, J., Horttanainen, M., Astrup, T. F., Saunila, M., & Happonen, A. (2022). Factors influencing household waste separation behavior: Cases of Russia and Finland. *Recycling, 7*(52), 1-15. https://doi.org/10.3390/recycling7040052

Appendix

Note. There is no need to get permission from the authors to use the scale.

Sayın öğretmen adayı,

Bu ankette size geri dönüşüme yönelik tutumlarınızı belirlemek üzere çeşitli sorular sorulmaktadır. Lütfen her cümleyi dikkatle okuduktan sonra, size uygun gelen seçeneği mutlaka işaretleyiniz. Unutmayın Doğru ya da Yanlış cevap yoktur. Sorulara içtenlikle cevap vermenizi rica ederiz. Katkılarınızdan dolayı teşekkür ederiz.

Demografik bilgiler **1.** Cinsiyetiniz: O Kadın O Erkek **2.** Sınıfınız: O 1. Sınıf O 2. Sınıf O 3. Sınıf O 4. sınıf	**Kesinlikle Katılmıyorum**	**Katılmıyorum**	**Biraz Katılıyorum**	**Katılıyorum**	**Tamamen Katılıyorum**
Sorumluluk ve Davranış					
1. Çevre dostu ürünler almaya özen gösteririm.	1	2	3	4	5
2. Yürüyüş yaparken gördüğüm çöpleri toplarım.	1	2	3	4	5
3. Ambalajların üzerinde çevre dostu ibaresi olup olmadığına dikkat ederim.	1	2	3	4	5
4. İnsanları geri dönüşüm yapmaya teşvik ederim.	1	2	3	4	5
5. Geri dönüşüm ve atık ile ilgili yerel veya küresel faaliyetlere aktif olarak katılırım.	1	2	3	4	5
6. Atıkları geri dönüşüm türlerine göre ayırırım.	1	2	3	4	5
7. Çevremdekilere örnek olmak için gördüğüm çöpleri toplayıp çöp kutusuna atarım.	1	2	3	4	5
8. Diğer insanların yere attığı çöpleri alıp çöp kutusuna atmam.*	1	2	3	4	5
9. Televizyon veya radyoda geri dönüşümle ilgili haberler ilgimi çeker.	1	2	3	4	5
10. Atıkları geri dönüşüm kutularına atıyorum.	1	2	3	4	5
11. Yere çöp atanları uyarıyorum.	1	2	3	4	5
12. Hangi atığın hangi renk atık kutusuna atılacağını biliyorum.	1	2	3	4	5
13. Çevremde gördüğüm atıkları çöpe atmaya çalışırım.	1	2	3	4	5
Bilinç ve Farkındalık					
14. Sokağa izmarit atmak çevresel bir problem oluşturmaz.*	1	2	3	4	5
15. Sıvı yağlar yemeklerde kullandıktan sonra zararsız hale gelir, bu nedenle geri dönüşüme gerek yoktur.*	1	2	3	4	5
16. Atık yağın lavaboya dökülmesinde bir sakınca yoktur.*	1	2	3	4	5
17. Geri dönüşüm belediyenin sorumluluğundadır. Bu nedenle atıkları geri dönüştürmek için çaba göstermemize gerek yoktur.*	1	2	3	4	5
18. Nehir, göl, deniz gibi su kaynaklarına atılan atıklar insanlar için tehlike oluşturmaz.*	1	2	3	4	5
19. Bir ürünün yeniden kullanımının çevresel bir faydası yoktur.*	1	2	3	4	5
20. Geri dönüşüm doğanın korunmasına yardımcı olur.	1	2	3	4	5
21. Geri dönüşümün öneminin farkındayım.	1	2	3	4	5
22. Biyoçeşitlilik geri dönüşümle korunabilir.	1	2	3	4	5
23. Sürdürülebilir kalkınma ve geri dönüşüm arasında bir ilişki yoktur.*	1	2	3	4	5
Ekonomik Değer					
24. Geri dönüşüm sadece pahalı ve değerli ürünler için yapılmalıdır.*	1	2	3	4	5
25. İnsanlar daha fazla gelir elde edebilmek için doğal kaynakları kullanabilir.*	1	2	3	4	5
26. Geri dönüşüm sürecinde yapılan masraflar yerine yeni bir ürün almak daha mantıklıdır.*	1	2	3	4	5
27. Geri dönüşümün ekonomik bir faydası yoktur. Bu nedenle geri dönüşüme gerek yoktur.*	1	2	3	4	5
28. Geri dönüşüm işlemi para israfına neden olur.*	1	2	3	4	5
29. Atıkların geri dönüşüm sürecindeki maliyetinden dolayı yerleşim yerlerinden uzağa atılmalıdır.*	1	2	3	4	5
30. Geri dönüşüm, kaynakların korunmasına yardımcı olur.	1	2	3	4	5
31. Geri dönüşüm maliyetli bir süreçtir.*	1	2	3	4	5
32. Eski bir ürünü hemen atıp yenisini alırım.*	1	2	3	4	5

* Olumsuz kodlanan ifadeler