

YIL / YEAR | 2025

CİLT - SAYI / VOLUME - ISSUE | 33 / 04

e-ISSN | 2147-9844

KASTAMONU ÜNİVERSİTESİ
KASTAMONU EĞİTİM DERGİSİ

KASTAMONU UNIVERSITY
KASTAMONU EDUCATION JOURNAL





KASTAMONU ÜNİVERSİTESİ KASTAMONU UNIVERSITY
KASTAMONU EĞİTİM DERGİSİ KASTAMONU EDUCATION JOURNAL

SAHİBİ | OWNER

Prof. Dr. Ahmet Hamdi TOPAL
Kastamonu Üniversitesi Rektörü | Rector of Kastamonu University

YAYINCI | PUBLISHER

Kastamonu Üniversitesi | Kastamonu University

BAŞ EDITÖR | EDITOR-IN-CHIEF

Doç. Dr. Muammer ERGÜN

YARDIMCI EDITÖRLER | ASSOCIATE EDITORS

Doç. Dr. Adem YILMAZ, Kastamonu Üniversitesi
Dr. İbrahim Fatih DEMİREL, Kastamonu Üniversitesi

TEKNİK EDITÖRLER | TECHNICAL EDITORS

Doç. Dr. Adem YILMAZ, Kastamonu Üniversitesi

DİL EDITÖRÜ | LANGUAGE EDITOR

Dr. Ersin ARSLAN, Kastamonu Üniversitesi

İLETİŞİM | CONTACT

Kastamonu Üniversitesi Eğitim Fakültesi, Kastamonu, TÜRKİYE

90 366 2803311

kefdergi@kastamonu.edu.tr

<https://dergipark.org.tr/tr/pub/kefdergi>



KASTAMONU ÜNİVERSİTESİ KASTAMONU UNIVERSITY
KASTAMONU EĞİTİM DERGİSİ KASTAMONU EDUCATION JOURNAL

EDİTÖR KURULU | EDITORIAL BOARD

Prof. Dr. Abdullah Çağrı BİBER, Kastamonu Üniversitesi

Doç. Dr. Emine KARASU AVCI, Kastamonu Üniversitesi

Doç. Dr. Erkan YANARATEŞ, Kastamonu Üniversitesi

Doç. Dr. Ezgi MOR, Kastamonu Üniversitesi

Doç. Dr. Gökhan KAYA, Kastamonu Üniversitesi

Prof. Dr. Hüseyin YOLCU, Kastamonu Üniversitesi

Doç. Dr. İbrahim KEPÇEOĞLU, Kastamonu Üniversitesi

Prof. Dr. İlhan ÖZGÜL, Kastamonu Üniversitesi

Dr. Öğr. Üyesi İsmail YILDIZ, Kastamonu Üniversitesi

Doç. Dr. Mehmet Koray SERİN, Kastamonu Üniversitesi

Doç. Dr. Melike FAİZ, Kastamonu Üniversitesi

Dr. Murat KONUK, Kastamonu Üniversitesi

Doç. Dr. Mustafa Öztürk AKCAOĞLU, Kastamonu Üniversitesi

Dr. Öğr. Üyesi Sema KARA, Kastamonu Üniversitesi

Doç. Dr. Sevcin CANDAN HELVACI, Kastamonu Üniversitesi



Kastamonu Eğitim Dergisi

Yıl:2025 Cilt:33 No:4

Kastamonu Education Journal

Year: 2025 Volume: 33 Issue:4

Kastamonu Eğitim Dergisi Uluslararası Bilimsel Hakemli
Dergidir.

Ocak-Nisan-Temmuz-Ekim Aylarında yılda 4 defa çevrimiçi olarak yayınlanır.

Taranmakta olduğumuz dizinler:

TR Dizin ULAKBİM		CrossRef	
H.W. Wilson Education Full Text		Google Scholar	
EBSCO Education Source		Index Copernicus	
Education Research Complete		ERIH PLUS	
SOBIAD		Cite Factor	
Türk Eğitim İndeksi		EBSCO- The Belt and Road Initiative Reference Source	

Kastamonu Üniversitesi
Kastamonu Eğitim Dergisi





| Research Article / Araştırma Makalesi |

Development of Perceived Physical Literacy Scale for All

Herkes İçin Algılanan Fiziksel Okuryazarlık Ölçeğinin Geliştirilmesi¹

Hüseyin SARIÇAM², Semih YILMAZ³

Keywords

1. Physical literacy
2. Public Recreation
3. Physical Activity
4. Recreation
5. Physical Education

Anahtar Kelimeler

1. Fiziksel okuryazarlık
2. Yerel Yönetimler
Rekreasyonu
3. Fiziksel Aktivite
4. Rekreasyon
5. Beden Eğitimi

Received/Başvuru Tarihi
27.01.2024

Accepted / Kabul Tarihi
10.10.2025

Abstract

Purpose: The aim of this study is to develop a scale for the perception of physical literacy on individuals participating in local government recreation activities.

Design/Methodology/Approach: In order to establish the construct validity of the scale, a model was first created through Exploratory Factor Analysis (EFA) with the first sample group; Then, Confirmatory Factor Analysis (CFA) was conducted with the data collected from the second sample group.

Findings: In EFA, sample adequacy was found to be .924 (>.60) and Bartlett's test of sphericity was significant at the $p < 0.01$ significance level. It is also seen that there are 3 factors with eigenvalues greater than 1, and the total variance they explain is 69.690%. A structure with a total of 19 items was obtained with a factor load value $> .45$. As a result of CFA, it is seen that the model is at a good level in all fit indices.

Highlights: The obtained scale has been shown to be a reliable and valid scale in determining the physical literacy people perceive in themselves.

Öz

Çalışmanın amacı: Yerel yönetimler rekreasyonu aktivitelerine katılanlar bireyler üzerinde fiziksel okuryazarlık algısına yönelik ölçek geliştirmek bu çalışmanın amacını oluşturmaktadır.

Materyal ve Yöntem: Ölçeğin yapı geçerliğinde öncelikle birinci örneklem grubuyla Açıklayıcı Faktör Analizi (AFA) yoluyla bir model oluşturulup; daha sonra ikinci örneklem grubundan toplanan verilerle Doğrulayıcı Faktör Analizi (DFA) yapılmıştır

Bulgular: AFA'da örneklem yeterliliği için .924 (>.60) ve Bartlett küresellik testi $p < 0.01$ önem düzeyinde anlamlı çıkmıştır. Ayrıca öz değeri 1'den büyük olan 3 faktörün olduğu görülmekte ve açıkladıkları toplam varyans %69.690'dir. faktör yük değeri $> .45$ olan toplam 19 maddelik bir yapı elde edilmiştir. DFA sonucunda uyum indislerinin tamamında modelin iyi düzeyde olduğu görülmektedir.

Önemli Vurgular: Elde edilen ölçek kişilerin kendilerinde algıladıkları fiziksel okuryazarlığı belirlemede güvenilir ve geçerli bir ölçek olduğu ortaya konulmuştur.

¹ This study comprises part of the first author's doctoral thesis.

² **Corresponded Author**, Marmara University, Institute of Health Sciences, Sports Management Science, Istanbul, TURKIYE; <https://orcid.org/0000-0001-5063-814X>

³ Marmara University, Faculty of Sport Science, Physical Education and Sport, Istanbul, TURKIYE; <https://orcid.org/0000-0001-6774-1047>

INTRODUCTION

Given the high prevalence of physical inactivity worldwide, introducing individuals to healthy, physically active lifestyles should be considered an important task of our modern societies (Carl et al., 2022). The increasing prevalence of sedentary living among children and adolescents has become an international concern, being responsible for the increased risk of childhood overweight and obesity, high blood pressure, and death (Li et al., 2021).

Understanding the importance of participating in physical activity, maintaining it throughout life, and knowing its positive effects on health is the key to physical literacy. When physical education is properly conducted by a trained teacher who follows the curriculum while providing physical activity with opportunities to learn, practice and participate, physical literacy development is inevitable (Hills et al., 2014). The concept of "physical literacy" is the individual's capacity to have an active lifestyle. Until recently, physical literacy studies were limited due to the lack of a common definition. The most relevant study of the concept of physical literacy belongs to Margaret Whitehead, who defined physical literacy as motivation, self-confidence, physical competence, knowledge and understanding to continue lifelong physical activity (Longmuir and Tremblay, 2016).

Physical literacy is a multifaceted term that encompasses areas such as movement competence, fun and enjoyment, self-confidence and social participation (Cairney et al., 2018). Physical literacy can be expressed as "the motivation, self-confidence, physical competence, knowledge and understanding required for an individual to continue lifelong physical activity at a certain level" (Whitehead, 2007).

Studies on physical literacy are aimed at comments and recommendations, and there is a lack of experimental studies. Due to the lack of studies and lack of evidence on physical literacy, it is not correct to directly associate it with lifelong physical activity (Longmuir and Tremblay, 2016). Since more experimental research is needed, a measurement tool to evaluate physical literacy is needed. In this study, a scale for the perception of physical literacy was developed for participants in local government recreation activities to fill the gap in the evaluation tool in the literature. This study aims to identify people's perceived physical literacy level inhibitors and to suggest solutions and intervene in their physical literacy journeys. Another aim of this study is to make positive contributions to individual and then public health through physical literacy awareness. For this reason, it gains importance as it will fill the gap in action-oriented research in the literature.

METHOD/MATERIALS

Study Group

The research was conducted in 2023-2024 with individuals benefiting from the sports centers within Istanbul Metropolitan Municipality Sports Istanbul facilities. In order to determine the construct validity of the scale, a model was first created through Exploratory Factor Analysis (EFA) with the first sample group; Then, Confirmatory Factor Analysis (CFA) was conducted with the data collected from the second sample group. For this purpose, the first study group in the research consisted of 242 people doing sports in Sports Istanbul facilities. 62% of the participants in this group were male (150), 38% were female (92); 38.8% had a bachelor's degree (94), 23.6% had a high school degree (57), 21.5% had an associate's degree (52), 8.3% had a master's degree (20), 4.1% had a secondary school (10), 2.0% had a doctorate (7), It was determined that 0.8% had primary school (2) education level. The lowest among the participants is 18 years old and the highest is 61 years old. The first place is 22 years old (31 people) with 12.8%, the second place is 28 years old (23 people) with 9.5%, and the third place is 18 years old (16 people) with 6.6%. There are 373 participants in the second study group. 61.9% of the participants in this group were male (231), 38.1% were female (142); It was determined that the largest percentage, 39.7%, was undergraduate (148), followed by 22.8%, high school (85), and 22.3%, associate degree (83). The age range of the participants varied between 18 and 61, and it was determined that 12.3% were 22 years old (46), 10.2 were 28 years old (38) and 7.2 were 18 years old (27).

Data Collection Process

Ethics committee permission for the research was received from Marmara University Health Sciences Institute Ethics Committee with date 20.06.2022 and protocol number 82. The "Google forms" survey tool of the www.google.com website was used in the research. The research data was filled in by the researcher to those who volunteered to participate and approved the consent form. Only demographic information was asked from the participants, but no identification information was requested.

Data Collection Tools

Development of the Measurement Tool

It was planned to develop the Perceived Physical Literacy Scale for All developed within the scope of the research in 5 stages, taking into account the scale development stages in the literature (creating the item pool, obtaining expert opinions, conducting a trial application, factor analysis, reliability analysis) (Metin et al., 2012).

Development of Item Pool

In developing the scale, the 18-item item pool of the Perceived Physical Literacy Scale developed by Sum et al. (2016) was used in the study, and 17 items were added by the researchers in line with the literature. According to DeVellis (2017), it is difficult to

determine the number of items that should be added to the item pool at the beginning of scale development studies. In order to increase internal consistency, a large item pool must be created when reaching the final scale. A large number of items helps the researcher to select items suitable for the purpose of the research. Widely used Likert-type scales measure thoughts, beliefs, and attitudes. (DeVellis, 2017). The reason why the prepared scale was designed as a Likert-type scale is to reflect the differences between ideas in the most accurate way. In this context, the draft scale was rated as a five-point Likert type consisting of the options "strongly disagree (1), disagree (2), Neutral (3), agree (4), strongly agree (5)".

Obtaining Expert Opinion

Obtaining expert opinion may be an indicator of whether the items in the scale adequately express what is intended to be measured. (Büyüköztürk et al., 2004). After ensuring content validity in the study, face validity was also ensured. Face validity refers to what a scale appears to measure rather than what it measures. (Oncu, 1994). The developed scale was presented to the opinion of a total of 4 expert academicians, including 2 sports scientists, 1 education scientist and 1 linguist, to ensure content and face validity. In order to make the scale suitable for the final version, spelling errors were examined in terms of meaning and suitability for purpose.

Pilot Study

The Perceived Physical Literacy Scale for All was administered to 15 participants who use the Spor İstanbul facilities. Feedback was obtained from the participants following the administration. Based on this feedback, no issues were identified regarding the comprehensibility of the items. Participants also indicated that there were no redundant items with the same meaning and that the scale was appropriate for the group to which it was administered. It was determined that the scale can be completed within 5–8 minutes.

Factor Analysis

The prepared draft scale was applied in accordance with the Marmara University Ethics Committee Decision taken within the scope of the research. It was examined whether the results obtained from the participants' answers exhibited normal distribution. For this purpose, the suitability of the data obtained from 242 participants for factor analysis and sample adequacy were examined in the first stage of the research. In order to determine the suitability of the data obtained from the perceived physical literacy scale measurement tool for factor analysis, anti-image, determinant coefficients, skewness and kurtosis coefficients were examined. In addition, the Kaiser Meyer Olkin (KMO) for sample adequacy and the Bartlett sphericity test for multiple normality were examined (Pallant, 2020). Then, Principal Component Analysis was used to extract factors in EFA; and the varimax technique, one of the orthogonal rotation techniques, was used to determine the rotation method of the factors (Can, 2017). In determining the removal status of the items in the scale, factor loading values of .45 or above were accepted as criteria (Büyüköztürk, 2009). In addition, the suitability of the factor structure resulting from the EFA was tested using Structural Equation Models. In this direction, the suitability of the factor loadings determined by the Exploratory Factor Analysis was determined by the Confirmatory Factor Analysis. In the Confirmatory Factor Analysis, RMSEA, CFI, GFI, AGFI, NNFI, SRMR and χ^2 (chi-square) criteria were evaluated as criteria to determine the suitability of the model. While the Exploratory Factor Analysis for the scale was tested with the SPSS 29.0.1 program, the AMOS program was used to apply the Confirmatory Factor Analysis.

Validity and Reliability Analysis

Validity is the level at which tests accurately measure the desired characteristics. There are three types: scope, construct and criterion validity. Scope validity is the adequacy of the items in the scale in terms of content and characteristics in order to measure the target behavior. Obtaining expert opinion is one of the methods used to ensure scope validity. A 70-80% agreement in the opinions expressed by experts and corrections made as a result of criticisms are sufficient to ensure scope validity. Construct validity is an indicator of how accurately an abstract concept is measured. Factor analyses, t-test scores, and test averages can be used for construct validity. Criterion validity is related to determining the relationship between test scores and the desired characteristic. Correlation coefficients and statistical significance results can be used to test criterion validity (Seçer, 2015). The consistency of all stages of the research from beginning to end is called reliability (Yıldırım and Şimşek, 2016). For this reason, the Cronbach alpha value is closely related to the number of items in the scale (Büyüköztürk, 2005). The Cronbach α internal consistency coefficient and the reliability coefficients of the sub-factors of the Perceived Physical Literacy Scale for All, which took its final form by performing factor analysis, were calculated. These values are included in the reliability analysis findings section. It is desired for this value to be over 0.70.

FINDINGS

In the study, exploratory and confirmatory factor analysis was conducted within the scope of the construct validity of the "Perceived Physical Literacy Scale for all.

Findings Regarding Exploratory Factor Analysis (EFA)

In the study, firstly, the suitability of the data obtained from 242 participants in the first application for factor analysis and sample adequacy were examined. In order to determine whether the data obtained from the perceived physical literacy scale for all measurement tool was suitable for factor analysis, anti-image, determinant coefficients, skewness and kurtosis coefficients were examined. In addition, the Kaiser Meyer Olkin (KMO) for sample adequacy and the Bartlett's sphericity test for multiple normality were examined and the test results are summarized in Table 1.

Table 1. Perceived physical literacy scale for all KMO and Barlet Test Result

Statistics	Value	
Kaiser-Meyer-Olkin Measure of Sampling Adequacy	.924	
Bartlett's Test of Sphericity	Approx. Chi-Square	3517.754
	df	171
	Sig.	.000

In Table 1, .924 (>.60) and Bartlett's sphericity test were found to be significant at $p < 0.01$ significance level. The findings show that the sample size is suitable for performing factor analysis. This also shows that the data come from a multivariate normal distribution (Kan and Akbaş, 2005). Later, in exploratory factor analysis, Principal Component Analysis was used to perform factor extraction, and the varimax technique, one of the orthogonal rotation techniques, was used to determine the rotation method of the factors (Can, 2017). In order to clarify whether the items in the scale should be removed or not, the factor loading values were accepted as .45 or above (Büyüköztürk, 2009). In addition, it was also examined whether the items had a loading value below a single factor. As a result of the factor analysis of the 19-item scale, it was determined that it explained 69.690% of the total variance and created a structure consisting of 3 factors. The findings regarding the analysis are presented in Table 2.

Table 2. Eigenvalues and Explained Variance of the Sub-dimensions of the Perceived Physical Literacy Scale for All

Component	Initials Eigenvalues			Rotation Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	9.568	50.359	50.359	6.014	31.650	31.650
2	2.008	10.570	60.928	3.985	20.971	52.621
3	1.665	8.762	69.690	3.243	17.069	69.690

When the exploratory factor analysis results are examined, it is seen that there are 3 factors with eigenvalues greater than 1 and the total variance they explain is 69.690%. The contribution of Factor 1 to the total variance is 31.650%, Factor 2 is 20.971% and Factor 3 is 17.069%. In social sciences, it is considered sufficient for the explained variance to be between 40% and 60% (Scherer et al., 1988). This result reflects that the total variance is quite sufficient.

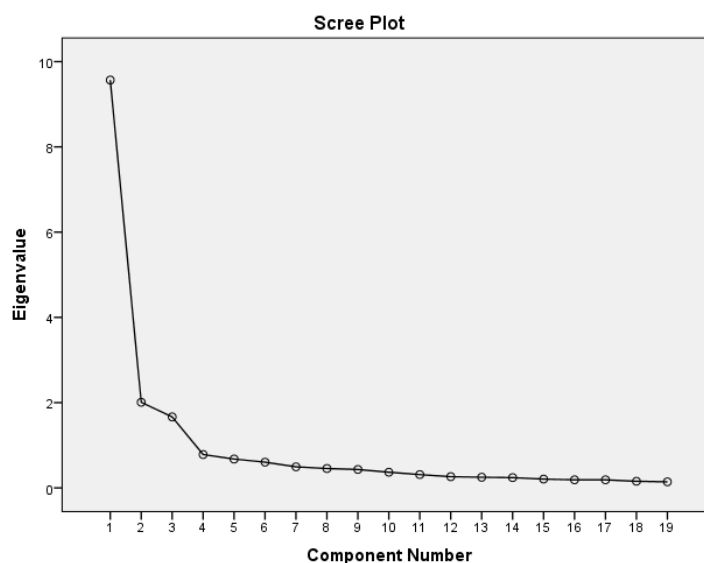


Figure 1. Scree Plot Test Result of Perceived Physical Literacy Scale for All

To verify the factor structure, a scree plot analysis was conducted. As shown in Figure 1, the elbow occurs after the third factor, indicating a three-factor solution. In addition, all items exhibit satisfactory structural integrity, characterized by substantial primary loadings on their target factors and an absence of substantively meaningful cross-loadings.

Table 3. EFA results for the Perceived Physical Literacy Scale for All

Item No	Self-Confidence and Sense of Self (Factor1)	Physical activity Skills and Sports Knowledge (Factor2)	Physical activity Attitudes (Factor3))	Item Quality (Çokluk et al. 2010)
1	.834			Excellent
2	.817			Excellent
3	.801			Excellent
4	.792			Excellent
5	.755			Excellent
6	.741			Excellent
7	.734			Excellent
8	.662			Great
9	.614			Good
10		.875		Excellent
11		.741		Excellent
12		.692		Great
13		.665		Great
14		.614		Good
15		.592		Good
16			.886	Excellent
17			.871	Excellent
18			.751	Excellent
19			.681	Great
Explained Variance (%)	31.650	20.971	17.069	%69.690

Table 3 shows the structure of the factor obtained after rotation and the factor loading values. While evaluating the findings in the table, attention was paid to the fact that the factor loading value was $>.45$ (Çokluk et al., 2016) and the difference between two factor loading values was at least $>.10$ (Büyüköztürk, 2009). As seen in the table, it was concluded that the factor loading values varied between .592 (Item 15) and .886 (Item 16). However, 16 items with a value below .45 and a loading value less than .10 among the items were removed from the scale and a structure consisting of a total of 19 items was obtained. These results show that the factor loadings obtained were high.

Table 4. Item Analysis of the Perceived Physical Literacy Scale for All

Item	Group	N	$\bar{X}$	SS	t	p	Item Total Correlation
1	Upper 27%	65	4.6462	.48188	14.943	.00	.788
	Lower 27%	65	2.9077	.80473			
2	Upper 27%	65	4.5538	.58712	15.844	.00	.731
	Lower 27%	65	2.7385	.71320			
3	Upper 27%	65	4.5538	.53124	12.320	.00	.636
	Lower 27%	65	2.8769	.96027			
4	Upper 27%	65	4.4615	.50240	13.087	.00	.767
	Lower 27%	65	2.9692	.76993			
5	Upper 27%	65	4.6923	.46513	12.669	.00	.713

Item	Group	N	$\bar{X}$	SS	t	p	Item Total Correlation
6	Lower 27%	65	3.3077	.74840	19.912	.00	.770
	Upper 27%	65	4.7538	.43412			
7	Lower 27%	65	2.7385	.69094	15.926	.00	.787
	Upper 27%	65	4.7385	.44289			
8	Lower 27%	65	2.8308	.85822	11.880	.00	.678
	Upper 27%	65	4.5538	.58712			
9	Lower 27%	65	3.2308	.67937	12.149	.00	.749
	Upper 27%	65	4.7231	.51562			
10	Lower 27%	65	3.1385	.91646	12.182	.00	.543
	Upper 27%	65	4.0923	.70096			
11	Lower 27%	65	2.4308	.84722	13.846	.00	.686
	Upper 27%	65	4.6615	.53843			
12	Lower 27%	65	2.8615	.89925	10.636	.00	.624
	Upper 27%	65	4.5538	.70779			
13	Lower 27%	65	3.0154	.92690	11.280	.00	.676
	Upper 27%	65	4.5692	.49904			
14	Lower 27%	65	3.3692	.69752	11.432	.00	.633
	Upper 27%	65	4.2923	.65486			
15	Lower 27%	65	2.9846	.64933	7.776	.00	.563
	Upper 27%	65	4.6462	.51329			
16	Lower 27%	65	3.8000	.71151	9.379	.00	.516
	Upper 27%	65	4.6923	.52806			
17	Lower 27%	65	3.4923	.88606	8.758	.00	.480
	Upper 27%	65	4.3538	.75892			
18	Lower 27%	65	3.0923	.87897	9.381	.00	.590
	Upper 27%	65	4.5538	.63813			
19	Lower 27%	65	3.2769	.89281	12.505	.00	.681
	Upper 27%	65	4.7231	.54508			
Reliability (α)= 0.94							

According to Table 4, it is seen that the corrected item total correlations of the items of the Perceived Physical Literacy scale for All vary between .788 - .480. This finding shows that the values are at the desired level (Büyüköztürk, 2004). On the other hand, the differences between the Lower and Upper 27% groups were found to be significant in the mean scores ($p<.01$). As a result of the findings, it can be said that the internal consistency is high in distinguishing people and items.

Findings Regarding Confirmatory Factor Analysis (CFA)

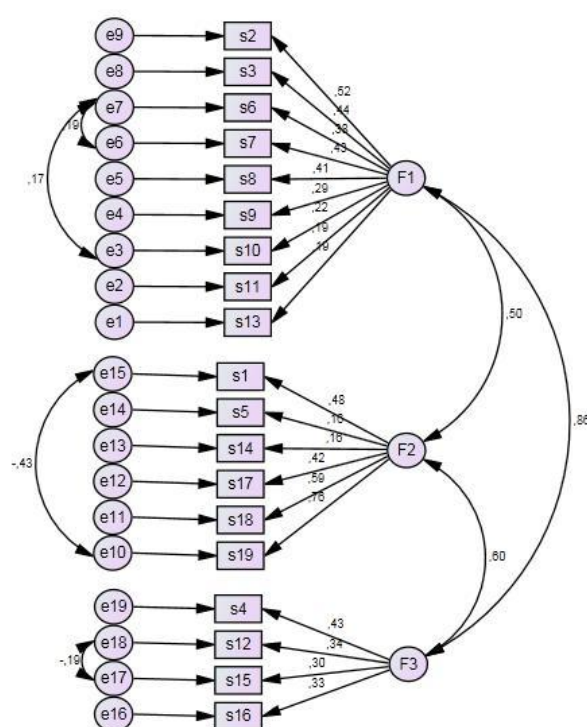
The aim is to determine the degree to which a structure previously determined by exploratory factor analysis is confirmed by the data collected later (Büyüköztürk et al., 2004). In the CFA (Confirmatory Factor Analysis) processes, the three-factor structure obtained from the EFA (Exploratory Factor Analysis) results was analyzed. The suitability of the factor structure was examined according to the goodness of fit and modification results. The goodness of fit of the HİAFOÖ scale calculated with CFA and the indexes accepted in the literature are given in Table 5.

Table 5. Confirmatory Factor Analysis Fit Indices

CFA Fit Index		Research Findings	Good Model Criteria (Çokluk, et al., 2010)
Chi-Square/Degrees of Freedom	χ^2/sd	2.65	<5
Goodness of Fit	GFI	.901	$\geq .85$
Adjusted Goodness of Fit	AGFI	.870	$\geq .85$
Comparative Fit Index	CFI	.96	$\geq .90$
(Non) Normed Fit Index	NNFI	.96	$\geq .90$
Standardized Root Mean Square Residual	SRMR	.714	$\leq .08$
Root Mean square Error of Approximation	RMSEA	.066	$\leq .08$

When the findings obtained in Table 5 as a result of CFA are compared with the values accepted in the relevant literature, it is seen that the model is at a good level in all fit indices. As a result of the findings, it was concluded that the model obtained is in compliance with the data.

Figure 2. Path Diagram of Perceived Physical Literacy Scale for All



Chi-Square= 382,3; df= 145, p-value=.0000, RMSEA=.066

When the CFA results were examined, the t values in the path diagram were checked for all items and it was determined that the lowest t value was 7.776 and had a significant t value at the .01 level. On the other hand, the images of the factor loading values were included in the path diagram shown in Figure 2. When the fit indices reported in Table 5 and the path diagram given in Figure 2 were examined, it was determined that the factor structure of this 3-factor scale was at a good level and the fit indices showed excellent and moderate model fit.

DISCUSSION

The Perceived Physical Literacy Scale for All developed within the scope of the research was planned to be developed in 5 stages, taking into account the scale development stages in the literature (creation of the item pool, obtaining expert opinions, conducting a trial application, factor analysis, reliability analysis) (Metin et al., 2012). In the development of the scale, the 18-item item pool of the Perceived Physical Literacy Scale developed by Sum et al. (2016) was used and 17 items were added by the researchers in line with the literature. The developed scale was presented to a total of 4 expert academicians, including 2 sports scientists, 1 educational scientist, and 1 linguist, in order to ensure content and face validity. The scale developed by Sum et al. (2016) consisted of 9 items and 3 sub-dimensions and was named under the titles of "Sense of Self-Confidence", "Knowledge and Understanding", and "Self-Expression and Communication". The study of Margaret Whitehead, who defined physical literacy as the motivation, confidence, physical competence, knowledge and understanding to maintain physical activity throughout life, is referred to as the study most associated with the concept of physical literacy (Longmuir and Tremblay, 2016).

Physical literacy is a multifaceted term that also encompasses areas such as movement competence, fun and enjoyment, self-confidence and social participation (Cairney et al., 2018). Physical literacy can be expressed as "the motivation, self-confidence, physical competence, knowledge and understanding required to maintain physical activity at a certain level throughout life as an individual" (Whitehead, 2007).

CONCLUSION AND RECOMMENDATIONS

In this study, which aimed to evaluate the factor structures of the Perceived Physical Literacy Scale for All, it was seen that the measurement tool consisted of 19 items and 3 sub-dimensions after the validity and reliability analyses were conducted and the factors; Items 1, 2, 3, 4, 5, 6, 7, 8 and 9 were named as "Self-Confidence and Self-Perception", items 10, 11, 12, 13, 14 and 15 were named as "Physical Activity Skills and Sports Knowledge" and items 16, 17, 18 and 19 were named as "Physical Activity Attitudes". Exploratory factor analysis (EFA) and confirmatory factor analysis (CFA) techniques were used in this study. The research was conducted with individuals benefiting from the sports centers within the Istanbul Metropolitan Municipality (IMM) Sports Istanbul in 2023-2024. In the construct validity of the scale, firstly, a model was created with the first sample group through Exploratory Factor Analysis (EFA); then, Confirmatory Factor Analysis (CFA) was conducted with the data collected from the second sample group. In this study, where the factor structure of the "Perceived Physical Literacy Scale for All" was evaluated, χ^2/df , RMSEA, SRMR, RMR, NNFI, CFI, GFI and AGFI fit indices were taken into account and it was found that all of them were at acceptable rates. Finally, as a result of the validity and reliability analyzes, Cronbach's alpha coefficient (α) = 0.94 was determined. Since the research findings were obtained only from the data collected from individuals who came to the sports centers belonging to local governments, it can be considered as a limitation of the research. As a result, it was revealed that the obtained scale is a reliable and valid scale in determining the physical literacy that people perceive in themselves.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, author-ship, and/or publication of this article.

Statements of publication ethics

We hereby declare that the study has not unethical issues and that research and publication ethics have been observed carefully.

Examples of author contribution statements

H.S and S.Y conceived of the presented idea. H.S developed the theory and performed the computations. H.S and S.Y verified the analytical methods. S.Y encouraged H.S to investigate and supervised the findings of this work. All authors discussed the results and contributed to the final manuscript.

Researchers' contribution rate

The study was conducted and reported with equal collaboration of the researchers.

Ethics Committee Approval Information

Ethics committee permission for the research was received from Marmara University Health Sciences Institute Ethics Committee with date 20.06.2022 and protocol number 82.

REFERENCES

- Bursal, M. (2019). SPSS ile temel veri analizleri. Anı Yayıncılık.
- Büyüköztürk, Ş., Akgün, Ö., Kahveci, Ö. & Demirel, F. (2004). The validity and reliability study of the Turkish version of the motivated strategies for learning questionnaire. *Educational Sciences: Theory & Practice*, 4(2), 207-239.
- Büyüköztürk, Ş. (2005). Anket geliştirme süreci. *Türk Eğitim Bilimleri Dergisi*, 3(2), 1-19.
- Büyüköztürk, Ş. (2009). Sosyal bilimler için veri analizi el kitabı: İstatistik, araştırma deseni, SPSS uygulamaları ve yorum (9. baskı). Pegem Akademi Yayıncılık.
- Cairney, J.; Veldhuizen, S.; Graham, J.D.; Rodriguez, C.; Bedard, C.; Bremer, E.; Kriellaars, D. A Construct Validation Study of PLAYfun. *Med. Sci. Sports Exerc.* 2018;50, 855–862.
- Can, A. (2017). SPSS ile bilimsel araştırma sürecinde nicel veri analizi (5. baskı). Pegem Akademi
- Carl, J., Barratt, J., Töpfer, C., Cairney, J., & Pfeifer, K. How are physical literacy interventions conceptualized?—A systematic review on intervention design and content. *Psychology of Sport and Exercise*, 2022;58, 102091.
- Çokluk, Ö., Şekercioğlu, G. & Büyüköztürk, Ş. (2010). Sosyal bilimler için çok değişkenli istatistik: SPSS ve LISREL uygulamaları. Pegem Akademi Yayıncılık
- De Vellis, R. F.(Ed.).(2017). Ölçek geliştirme kuram ve uygulamalar.Nobel.
- Kan, A. & Akbaş, A. (2005). A study of developing an attitude scale towards chemistry. *Mersin University Journal of the Faculty of Education*, 1(2), 227-237.
- Hills, A. P., Dengel, D. R., & Lubans, D. R. Supporting public health priorities: Recommendations for physical education and physical activity promotion in schools. *Progress in Cardiovascular Diseases*, 2014;57, 368–374. <http://doi.org/10.1016/j.pcad.2014.09.010>
- Li, M. H., Sum, R. K. W., Sit, C. H. P., Liu Y. ve Li, R. Perceived and actual physical literacy and physical activity: A test of reverse pathway among Hong Kong children. *Journal of Exercise Science & Fitness*, 2021; 19(3), 171-177.
- Longmuir, P. E., & Tremblay, M. S. Top 10 research questions related to physical literacy. *Research quarterly for exercise and sport*, 2016;87(1), 28-35.
- Metin, M., Birişçi, S., Coşkun, K. & Kolomuç, A. (2012) A study on developing “basic computer use performance scale (BCUPS)” for primary students, *Procedia Social and Behavioral Sciences*, 46, 1771 –1775
- Öncü H. (1994). Eğitimde Ölçme ve Değerlendirme. Matser Basım San. Ve Tic. Ltd. Şti.
- Pallant, J. (2020). Spss Survival Manual: A step by step guide to data analysis using ibm spss. Routledge. (Çev.Balcı S. &Ahi B.)
- Scherer, R. F., Luther, D. C., Wiebe, F. A. & Adams, J. S. (1988). Dimensionality of coping: Factor stability using the ways of coping questionnaire. *Psychological Reports*, 62(3), 763-770. <https://doi.org/10.2466/pr0.1988.62.3.763>
- Seçer, İ. (2015). Spss ve Lisrel ile pratik veri analizi (Genişletilmiş 2. Baskı).Anı Yayıncılık.
- Sum, R. K. W., Ha, A. S. C., Cheng, C. F., Chung, P. K., Yiu K. T. C., Kuo, C. C., Yu, C. K., & Wang, F. J. (2016) Construction and validation of a perceived physical literacy instrument for physical education teachers. *PLoS ONE*, 11(5), 1-10.
- Whitehead, M. Physical literacy: Philosophical considerations in relation to developing a sense of self, universality and propositional knowledge. *Sport, Ethics and Philosophy*, 2007; 1(3), 281-298.
- Yıldırım, A. & Şimşek, H. (2016). Sosyal bilimlerde nitel araştırma yöntemleri(10. Baskı).Seçkin Yayıncılık.

Appendix.1

Perceived Physical Literacy scale for All (Herkes İçin Algılanan Fiziksel Okuryazarlık Ölçeği)	Kesinlikle Katılmıyorum	Katılmıyorum	Kararsızım	Katılıyorum	Kesinlikle Katılıyorum
1. Yaşıma göre fiziksel olarak iyi görünüme sahibim.	1	2	3	4	5
2. Spora karşı pozitif tutum ve ilgiye sahibim.	1	2	3	4	5
3. Sağlıklı ve zinde kalmak için kendimi yönetme becerilerine sahibim.	1	2	3	4	5
4. Güçlü iletişim becerilerine sahibim.	1	2	3	4	5
5. Zorlu doğa koşullarında fiziksel aktiviteyi sürdürebilirim	1	2	3	4	5
6. Spor yapmayı bir alışkanlık haline getirebilirim.	1	2	3	4	5
7. Spor sayesinde arkadaşlıklar kurarım.	1	2	3	4	5
8. Güncel spor trendlerini öğrenmeye hevesliyim.	1	2	3	4	5
9. Kendimi spor yapmak için motive etmenin yolunu bulurum.	1	2	3	4	5
10. Yeterli temel hareket (Çömelme, sıçrama, koşma, atlama vb.) becerilerine sahibim.	1	2	3	4	5
11. Beden eğitimi bilgilerimi yaşam boyu uygulayabilirim.	1	2	3	4	5
12. Performansımı ve hareket yeteneğimi nasıl geliştireceğimi bilirim.	1	2	3	4	5
13. Çeşitli fiziksel aktivitelerde iyi performans gösterebilirim.	1	2	3	4	5
14. Çeşitli spor branşları hakkında başkalarıyla tartışacak düzeyde bilgi sahibiyim.	1	2	3	4	5
15. Çeşitli spor branşları hakkında başkalarıyla tartışacak düzeyde bilgi sahibiyim.	1	2	3	4	5
16. Kendimin ve diğer insanların spor yapmasından memnuniyet duyarım.	1	2	3	4	5
17. Fiziksel aktiviteye katılmaya hevesliyim.	1	2	3	4	5
18. Fiziksel aktivite sağlık gelişimim için önemli değere sahiptir.	1	2	3	4	5
19. Fiziksel aktivite sosyal etkileşim kurmamda önemli yer tutar.	1	2	3	4	5



| Research Article / Araştırma Makalesi |

What Happened at Home During the Emergency Remote Teaching? A Phenomenological Study on Experiences of Turkish Parents

Acil Uzaktan Eğitim Sırasında Evde Neler Oldu? Türk Ebeveynlerin Deneyimleri Üzerine Fenomenolojik Bir Çalışma

Murat Sümer¹, Nesime Can Angın²

Keywords

Family relations
Emergency remote teaching
COVID-19
Primary school students
Phenomenology

Anahtar Kelimeler

Aile ilişkileri
Acil uzaktan eğitim
COVID-19
İlkokul öğrencileri
Fenomenoloji

Received/Başvuru Tarihi
21.04.2025

Accepted / Kabul Tarihi
10.10.2025

Abstract

Due to the COVID-19 pandemic, schools were forced to suspend all face-to-face classes around the world and education occurred at home. Millions of students and their parents had to adapt to 'emergency remote teaching' (ERT) without any prior knowledge or preparation. This study, using a qualitative method with a phenomenological approach, was designed to understand parents of primary school students' experiences during the ERT during the COVID-19. The data were collected through 15 semi-structured interviews with parents in Turkey. After content analysis, four major themes emerged: (a) parent experiences, (b) problems, (c) support, and (d) family relations. The results indicated that parents experienced various changes in their routines at home, roles and responsibilities related to their children's education, and family relations.

Öz

COVID-19 salgını nedeniyle, okullar tüm dünyada yüz yüze eğitime ara vermek zorunda kalmış ve eğitim evde devam etmiştir. Bunun sonucunda da, milyonlarca öğrenci ve ebeveynleri, herhangi bir ön bilgi veya hazırlık olmaksızın acil uzaktan eğitime uyum sağlamak zorunda kalmıştır. Bu çalışma, ilkökul öğrencilerinin ebeveynlerinin COVID-19 dönemindeki deneyimlerini anlamak üzere nitel bir çalışma olarak tasarlanmıştır. Veriler, Türkiye'deki ebeveynlerle yapılan yarı yapılandırılmış görüşme yoluyla toplanmıştır. İçerik analizi sonucunda dört ana tema ortaya çıkmıştır; (a) ebeveynlerin deneyimleri, (b) sorunlar, (c) destek ve (d) aile ilişkileri. Sonuçlar, ebeveynlerin evdeki rutinlerinde, çocuklarının eğitimiyle ilgili rol ve sorumluluklarında ve aile ile ilişkilerinde çeşitli değişiklikler yaşadıklarını göstermektedir.

¹ Corresponded Author, Uşak University, Faculty of Education, Department of Instructional Technologies, Uşak, TÜRKİYE; <https://orcid.org/0000-0003-1566-4355>

² Ankara University, Faculty of Educational Sciences, Department of Guidance and Psychological Counselling, Ankara, TÜRKİYE; <https://orcid.org/0000-0002-6448-1275>

INTRODUCTION

Coronavirus Disease 2019 (COVID-19) caused a global crisis and forced approximately 1.5 billion students to change their formal education (UNESCO, 2020). According to Williamson et al. (2020), this pandemic has caused a serious transformation process in social, cultural, economic, and other societal areas, in addition to being a serious health crisis. Global education systems have also had to transform, and schools have been among the institutions most affected by this process. These unanticipated school closures affected all stakeholders in teaching and learning, including students and their parents. Because of the pandemic, schools in Turkey were also forced to suspend all face-to-face classes and transition to distance learning to continue teaching and learning.

Aydin (2011) defines distance learning as a learning process in which learners are separated from one another but continue to interact and learn through information and communication technologies. According to Bozkurt and Sharma, distance learning gives students autonomy, responsibility, flexibility, and choice. What is being done during the pandemic is restricted to delivering content to remote learners. Therefore, distance learning and the educational practices during the pandemic should be separated from each other. Educational practice during the pandemic is referred to as emergency remote teaching and it is defined as a method of instructional delivery by educational institutions (Hodges et al., 2020). As a result, it should be regarded as a temporary solution to the pandemic, as opposed to distance learning. The transition to emergency remote teaching has introduced new challenges at a social and technological level, influencing both children's and parents' physical and mental health (Misirli & Ergulec, 2021).

Parents' Experiences during the Emergency Remote Teaching

When parents accompany their children, they become one of the most important stakeholders of teaching and learning during emergency remote teaching (Misirli & Ergulec, 2021). Parents became teachers and schools expanded into homes. According to Dong et al. (2020), during the emergency remote teaching, parents' beliefs and attitudes toward distance learning were found to be controversial. On the one hand, parents have begun recognizing the value of digital devices and have become more comfortable with their children's use at home, as well as supporting the appropriate use of digital devices for learning. Parents, on the other hand, were concerned about the dangers of unrestricted digital use, as well as the impact of digital use on children's social and physiological development (Plowman et al., 2012). According to Dong et al. (2020), parents with young children in China had fewer positive beliefs about the value of distance learning. Only a small proportion of participants believed that distance learning has better learning content (18.4 %), better learning outcomes (11.0 %), and is more efficient (12.6 %) than traditional methods. Furthermore, the parents believed that traditional methods were superior to distance learning in terms of creating a learning atmosphere that resulted in better learning outcomes. Weaver and Swank (2021) stated that parents believed that some students struggled with distance learning. Also, some parents expressed their concerns that teachers assigned "busy work" in distance learning, which, according to studies, can be damaging to students' academic performance (Snelling & Fingal, 2020).

Additionally, a reason why the parents have unfavorable beliefs toward distance learning is their changing routines. Emergency remote teaching also changed parents' daily routine, which is important to establish a sense of normalcy and foster a positive learning environment (Weaver & Swank, 2021). Nearly 90% of parents reported that their lives had changed since the COVID-19 pandemic, with 44% reporting a significant change (Pew Research Center, 2020). When parents are forced to fill their children's time at home, without contact with peers or teachers, it creates a slew of surprises, tensions, and conflicts, necessitating a different organization of daily activities and frequently making everyday life difficult (Parczewska, 2020). During the pandemic, children were required to study at home in a new routine, and parents spent more time with their children. Sometimes, parents become teachers. According to Trzcinska-Krol (2020), most parents exert control over their children's work by monitoring their children's learning, homework, and test taking. When the children's performance fell short of their parents' expectations, it led to disagreements and a negative impact on the parents' attitudes toward distance learning (Wu et al., 2020).

Last but not least, during the pandemic, parents experienced increased levels of mental distress, according to Patrick et al. (2020). Since March 2020, 27% of parents have reported worsening mental health and 14% have reported worsening behavioral health in their children. Depression, anxiety, and stress were significantly higher in parents (Wu et al., 2020). Because of these problems, during the emergency remote teaching, many countries have increased support to parents, including the establishment of a dedicated educational TV channel for parents on how to be teachers at home, particularly for early childhood education (UNESCO, 2020).

To summarize, it can be concluded from the literature that students and their parents may not have enough time to prepare or receive adequate support for emergency remote teaching. Despite the fact that there is parental support available, we know little about the transition process of parents to emergency remote teaching, what issues they faced, how they were supported, and how they overcame the problems they faced. As a result, this study investigates ERT experiences from the perspective of parents in Turkey during the COVID-19.

Context of the Study

Turkey, as many other countries, followed strict protocols and shut down schools. There were approximately 18 million students and 1 million teachers in compulsory education levels in Turkey. Due to the COVID-19 pandemic, the Ministry of National Education (MoNE) of Turkey also decided to transition to emergency remote teaching in March 2020. Services for elementary,

secondary, high school, and special education students were offered through the Education Information Network (EBA), which is a distance learning platform that can be accessed through the internet, as part of the MoNE's measures. Furthermore, the MoNE developed a free television channel called EBA TV to provide lessons within the framework of the daily curriculum to elementary, secondary, and high school students (Mengi & Alpdogan, 2020).

Turkey is regarded as a developing nation. According to Gürmen and Kılıç (2022), women's employment in Turkey was 25.3% in 2002 while was 28.2% in 2021. In 2016, 70.7% of women stated that being preoccupied with housework was a reason for not working. However, the family structure in Turkey today is changing as women gain economic independence and move away from the traditional caregiver role.

Although all students needed support during the transition, specifically elementary school students and their parents in Turkey may need extra help due to children's developmental stage and lack of school experiences. As a result, we saw the need to investigate ERT through the experiences of families with primary school-aged children throughout the pandemic and to explain the importance of emergency remote teaching during COVID-19. Knowing about the experiences of parents who have had ERT can be beneficial for future pandemic-like events in terms of family relations and educational effectiveness, and it can supplement the limited literature.

METHOD

This study describes the essence of ERT during the COVID-19 pandemic by exploring ERT from the lived experiences of parents who have children at elementary school in Turkey during the 2020–2021 school year. For this purpose, answers to the following research questions were sought:

1. What were the experiences of parents during ERT?
2. How did parents deal with the issues they experienced during the ERT?
3. How was the family relation affected by the emergency remote teaching?

To answer these questions, this study is designed as a phenomenology, which is one of the qualitative methods in education. Phenomenology is an interpretive methodology that focuses on accessing participants lived experiences by using semi-structured interviews (Patton; 2015; Van der Mescht, 2004). In other words, phenomenology is a qualitative method of investigation in which researchers attempt to discover the meaning of human beings' lived experiences as they exist in the world (Chamberlain, 2009).

The study of consciousness and the content of conscious experience, such as judgements, perceptions, and emotions, is known as phenomenology (Balls, 2009). Researchers conducting phenomenology want to know what it was like to live the experience, not just how the person reacted to it (Moustakas, 1994; Munhall, 2007). Therefore, this is the strengthening of phenomenology, and it is a good fit for this study, as this study describes the essence of ERT during the COVID-19 pandemic by exploring ERT from the lived experiences of parents who have children at elementary school in Turkey.

Participants and Data Collection

In phenomenology, participants are purposely selected based on the basis of experience of the phenomenon under investigation (Van der Mescht, 2004). Participants of this study were parents with children in the elementary school and experienced emergency remote teaching at least one semester. When recruiting participants for the research, the purpose was to choose parents with at least one kid at the elementary school and participate in emergency remote teaching at home. We interviewed 15 parents whose children were in elementary school during the 2020–2021 school year. Participants are between the ages of 35–43, middle-income group, and college graduates. We only included elementary school students' parents in the study because elementary education plays a critical role in people's lives as one learns reading, writing, and arithmetic, and enhances social emotional development at this level. Therefore, parents' experiences and views on ERT became significant during the pandemic.

Before the study, an application for research permission was submitted to and approved by the institutional review board. Following that, an invitation link was created through Google forms, including the purpose of the study, and the link was distributed in various social media platforms. Parents who were interested in participating in this study filled out the form by indicating their interview preference and availability. Based on their preference and availability, the data were collected with semi-structured interviews via video or phone calls. We did not receive a large number of responses, and some of them later changed their minds and did not participate in the study. As a result, we contacted the majority of the parents who responded.

Interview protocol was developed prior to the interviews and was used during the interviews. According to the protocol, interviews began with the main research questions and continued with related questions based on participant responses. Both researchers joined the interviews. Before the interview, participants' consent was obtained, and the purpose of the study was explained. It was also highlighted that participation was entirely voluntary and that they could prefer out of the study at any time. They took about 20 minutes on average.

Reliability of the Study

Both the researchers took part in conducting interviews, transcription, and coding processes. The data were analyzed via content analysis. The validity and reliability of qualitative research can be ensured using a variety of approaches (Creswell & Poth, 2016; Golafshani, 2003), and researcher triangulation is one method (Denzin, 2017), which refers to the use of multiple researchers from different backgrounds. To accomplish this, one researcher with a background in online education and the other in counseling categorized the interview transcripts and created themes on their own. Finally, they collaborated to develop shared themes.

FINDINGS

The current study sought to investigate parents' experiences with emergency remote teaching (ERT) during the COVID-19 pandemic. After the analysis, the following four major final themes emerged: (a) parent experience, (b) problems, (c) support, and (d) family relations. Figure 1 depicts the sub-categories for each main theme. The findings on these themes are given below with verbatim quotes of the participants.

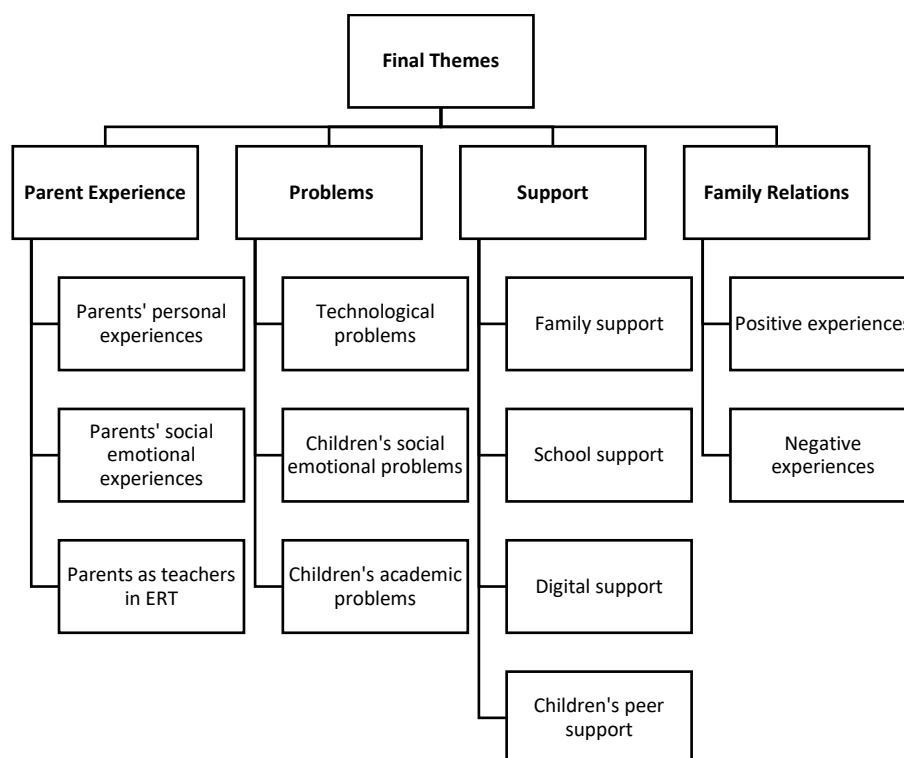


Figure 1: Themes

Parent Experiences

COVID-19 pandemic and ERT led to radical changes in parents' lives, and many parents and children experienced unique challenges with drastic changes. Our participants shared their experiences of how ERT impacted their own lives as well as their children's lives at home. The first theme, parent experience, had three categories: parents' personal experiences, social emotional experiences, and parents as teachers in ERT.

In terms of parents' personal experiences, all participants shared that their personal lives transformed in one way or another. Due to ERT, participants reported having a busy schedule with their children and needed to develop a new routine at home. For instance, P14 stated that "We tried to be careful about carrying our routine at home including getting up early in the morning, having our breakfast, and then going in front of the screen for classes." All participants except P3 shared that their work situation impacted how they reorganized their new routine. Due to the COVID-19 pandemic, a radical change occurred in how much people work from home. P9 described that their routine included a balance between ERT and work requirements, and stated "We wake up at the same time in the morning and after we have our breakfast, the day starts. My daughter's online lesson starts. I mean, her class starts at 9:00 am. After that, the classes continue until 11:30. We are working on the same desk. While my daughter attends [her] classes, I work on my job with her. Then, we have lunch and later we have dinner...This is almost a routine for us"

Participants also expressed their emotions and how tiring this period was. Ten participants (P1, P2 P3, P6, P8, P10, P11, P12, P13, and P14) shared their conflicted feelings about being extra busy and feeling safe for their family. For instance, P13 stated that "So, it was more tiring for us. You go home; you must work at home... We prioritize our children, so that their homework needs to be finished first, and then we must run errands." P6 also mentioned that "...Even when I was at work, I was constantly checking the time and my children's class schedules. I was worried that they would miss their online classes." Additionally, P12 reported

having mental health problems and stated that “I work in a high school and teach classes for 26 hours a week. As a result of teaching my own classes and keeping track of my children’s classes, my mental health began to deteriorate.” However, P13 shared positive emotions about the change as their children were at home due to ERT and not facing the risk of having COVID-19 by stating “Since our children did not go to school during the pandemic, we felt better, frankly, there was no fear of meeting a sick child or a teacher.”

Nine participants (P1, P2, P3, P4, P8, P10, P11, P14, and P15) shared that due to the ERT, parents owned new responsibilities such as being a teacher at home; because they feel responsible for how much their children learned. P3 shared that “Some of our roles have changed. The parents became teachers” P2 supported this notion and shared that “Sometimes he gets distracted and misses what the teacher is saying, then he does not understand the subject. ... When he does not understand, the burden is on us. We are trying to explain.” Parents also felt responsible for helping children with attending classes and completing homework on time. P10 reported that “I feel like it’s up to me to teach. I need to understand the homework, check the homework, and help with the parts that my child did not understand. These are all on me because of remote learning.”

Problems

Transition to mandatory ERT may be uneasy for many individuals. Considering the pandemic - related problems such as lockdowns and social isolation, the transition became more problematic for many families. In our study, participants reported their ERT - related problems under three categories: technological problems, children’s social emotional problems, and children’s academic problems.

Twelve participants (except P1, P5 and P12) reported experiencing technological problems during the transition to ERT. Participants reported using mobile data to connect to the Internet at home before ERT; however, their mobile data was not enough for their children and Wi-Fi connection became a necessity. P8 reported that “At first, she was attending live classes one day a week. At that time, we were using mobile internet data, which was enough. But then, we had to purchase Wi-Fi at home and a new tablet.” In addition to internet expenses, some participants had to purchase a new tablet or laptop for their children. For instance, P2 shared that “When the ERT first started during [my child’s] first grade, we did not have technological equipment at home. She did not have a tablet, nor did we have the Internet [WIFI] at home. Then, we bought a tablet, and then got the Internet.” After purchasing Wi-Fi and tablets, a few parents still experienced internet connection problems. P14 shared their internet -related problems and stated that “The most challenging part was for me to keep track of her class schedule and control whether she attended or left the class on time without having any problems. I constantly worry about her losing the internet connection and not being able to connect back.” P7 also agreed and shared that “First, we had an internet connection problem. After overcoming the connection problem, we needed to learn the interface of the program [for ERT].” P7, P15 also experienced problems related to online participation in the classes and mentioned “Our major problem was entering the password. The system was asking for a password to connect Zoom, and of course, our child could not enter it.... I had to go back and forth to enter the password for her.”

Almost all participants (except P2, P7, and P15) shared their observations about their children’s social emotional problems during ERT. During the COVID-19 pandemic, individuals are likely to worry about the risk of contracting COVID-19; therefore, they may not allow their children to go out, even if their children need to be active developmentally. For example, P11 mentioned that “Our children were stuck at home during this period...Their desire to play with their friends was being precluded and they were angry...This was a very problematic time in terms of socialization.” Similar to P11, P6 also shared that “Children are already bored because they stayed at home all the time. If the schools were open, they would go out and see their friends. So, it has been pretty tough right now.” Additionally, due to limited socialization, the time that parents monitored their child’s use of digital media was increased. P8 stated that “I should say that [my kid’s] screen time [usage] is increasing in the first place. I am not sure how the increased screen time will affect our child in the future.” P3 also mentioned increased screen time and socialization problems, stating “My child missed his friends so much. He had to spend a lot of time inside the house. Also, it is getting harder and harder to get the phone away from my son.”

Thirteen participants (except P3 and P11) shared their children’s academic problems during ERT. Parents of first-grade students experienced extra difficulty during the adaptation to ERT. P1 shared that “We had a hard time at first, and frankly, she had a hard time adapting to ERT because she is too young. She had difficulty communicating with the teacher through a computer or phone. She also had difficulty doing what he [teacher] said and following the instructions given.” Some other academic problems appeared as a lack of motivation, lack of focus, and learning loss. P8 described the problem as “The problem is the lack of motivation. They [his children] also had trouble staying motivated during online classes.” P6 also reported and shared that “It is very difficult for children to sit in front of the computer or the phone all the time.” When children did not understand the topic, parents looked for ways to help their children academically. P10 shared that “I could not support my child as I wish to and I felt inadequate. Since I felt tired, we arranged private teaching lessons [for our child].”

Supports

Although parents encountered numerous problems during the ERT, our participants reported that the support they received from various sources made the process easier. In this study, participants reported four main sources of support that are family support, school support, digital support, and children’s peer support.

In terms of family support, twelve participants (except P8, P10 and P11) reported receiving support from both their immediate and extended family. P3 highlighted the support she received from her husband and older daughter, stating "Of course, I received some help from my husband.... Also, my daughter helps me when I run out of patience. In that case, she immediately steps in and takes it over. ... My husband may say, let me take it over or I may say that to him." P6 also received help from her older child and stated that "My eldest son helped me a lot in this process. He has been a great help in keeping track of both his own classes and his brother's classes. We were also in constant dialog with the teachers." P1, P2, and P7 stated the significant support they received from their extended family during ERT. For instance, P2, who is a single mother living with her parents shared that "Everyone helps my daughter in our household. Let's say she has a class and we need to go somewhere. We make plans based on my daughter's class schedule."

Several participants have highlighted the importance of school support. Eight participants (P2, P6, P8, P9, P10, P12, P14 and P15) reported needing support from teachers and school counselors with online classes, assignments and general problems related to ERT. For instance, P8 shared that "We asked the teacher what we can and cannot do about our problems. Then, we tried to get help from the school counselor. We asked the counselor what we can do to make our situation better." P9 also shared receiving help both from the teacher and school counselor by stating "We sometimes had problems with homework and asked for help from our teacher. ...Also, the school counseling center provided support and the school counselor listened to our child's problems."

P7, P8, and P12 emphasized the importance of digital support during ERT. Both schools and parents had to provide technology-related support for children. For instance, P7 mentioned that "Currently, I can only provide technical support due to my location. I'm away from home but provide technical support... For example, sometimes there is a problem with the computer or the internet. I can reach her computer online and fix the problem."

In terms of parents' understanding of their children's peer support, P7, P9, and P15 shared that they assisted their children to find a way to keep them interacting with their peers. For instance, P7 arranged online meetings for their child and friends. Additionally, P9 communicated with other parents to arrange regular park gatherings for their children; "To cope with difficulties, we have tried several things. For instance, we talked to a few parents of our daughter's close friends and decided to get them together once a week in a park. So, they could socialize a little bit, but after the curfews, we did not continue."

Family Relations

Emergency remote teaching during the COVID-19 pandemic impacted various areas, including family relations. Almost all of our participants stated that their family relations were impacted both positively and negatively during this time. In terms of positive experiences, ten participants (P1, P2, P3, P4, P5, P7, P8, P10, P11 and P14) shared that their family relations strengthened in this process because of spending more time together. P11 reported that "We were delighted because we started to spend more time together as a family at home. So, we liked it. Her father was also at home. We noticed that we miss each other and fulfilled our longing." P1, a teacher, also reported that "Since we are always at home, we are in constant communication with each other. Previously, when we went to school, we were disconnected at home as we were both tired after school." Additionally, P5 shared that she fulfilled her longing toward her children during this time as she had to go back to work right after her maternity leave by stating "I have worked for the government for 25 years. I have not been did to spend much quality time with my children after my maternity leave. So, I'm happy with being at home with my family."

In terms of negative experiences, ten participants (P2, P3, P4, P8, P9, P10, P11, P12, P13 and P14) reported that they argued with their kids and tension raised from time to time. P9 shared an example and stated that "Problems arose sometimes and there were points where the limits of my tolerance were pushed." Additionally, P10 shared that "My relationship with my child has deteriorated...Of course, there is no problem in our bond of love, maybe we are even more connected, but I think we are worn out at certain points."

DISCUSSION

Due to the COVID-19 pandemic, governments took precaution of closing schools in many countries. In Turkey, education in primary, secondary, and high schools were all suspended in March 2020 (UNESCO, 2022), and ERT had started without meticulous preparation for over 5 million primary school students (MoNE, 2021). This study focused on parents' experiences during ERT in Turkey, because primary education is significant for children, as they learn reading, writing, and arithmetic as well as enhance their social emotional skills in those years. Therefore, we designed a study to explore ERT from the perspectives of parents who have children at the primary school level in Turkey during 2020–2021 academic year. After analysis, four main themes emerged (a) parent experience, (b) problems, (c) support, and (d) family relations to answer three main research questions.

With the first research question, we aimed to explore participants' personal experiences during ERT and study participants shared how their personal and emotional experiences were transformed. Due to nationwide lockdown, working from home, and ERT, parents' responsibilities were drastically increased. During the pandemic, children had to stay home and participate in ERT to learn and enhance their skills under their parents' supervision. In this period, participants of this study reported that, as primary caregivers, they felt responsible for their children's academic, social, and emotional developments while keeping their work-related responsibilities. In other words, participants reported that they owned a teacher role to support their children to participate in live/online class sessions and to complete their homework. Misirli and Ergulec (2021) reported that approximately

half of the parents assisted their children with technology and then left them alone during the live class sessions. During the live sessions, 34% of the parents either stood by their children or actively participated in the remote teaching. Our analysis of the interviews contributes to these findings in a way that parents responsibilities were increased. Under such circumstances, parents may need to be prepared for extra obligations, and as a result as Patrick et al. (2020) suggested parents may experience increased levels of mental distress.

We also aimed to explore what kind of problems participants encountered and how they dealt with such issues during the ERT. With the analysis of the interviews, we found that participants experienced technological problems and had to purchase WIFI, tablet or laptop. This might cause financial problems for some families. Additionally, individuals who have limited knowledge on use of tools such as basic programs and internet use could also create issues. Therefore, it is important to consider socioeconomic factors and technological needs of participants in future ERT related studies. Study participants also stated problems related their children's social, emotional, and academic problems. As their children stayed at home and participated class sessions online, parents were able to closely monitor their children's needs in various areas. Similarly, researchers found that parents sharing their children's social needs were somewhat addressed during the COVID-19 pandemic (Misirli & Ergulec, 2021). Furthermore, due to limited socialization, several parents reported that screen time usage of their children increased exceedingly. Likewise, Plowman and colleagues (2012) also stated that parents concerned about the dangers of unrestricted digital use. Participants additionally reported that their children experienced academic problems including attention and focus problems. According to study participants, some children had difficulty staying in front of a screen during classes, and others had difficulty understanding the subject and homework. Although with ERT, children could continue their education during the pandemic, based on our interviews' findings, primary school students were too young to self-regulate. Therefore, parents might need some extra guidance in terms of how to regulate some of the expected problems when they arise. School counselors and educators may address some of these needs arranging online meetings with parents, and creating informative digital materials to distribute. Also, educating parents about how to balance children's screen time and outdoor time by providing practical solutions can be helpful.

In the second research question, we also asked study participants about how they dealt with problems during ERT and our analysis results of the interviews suggested that receiving and providing support was essential ways to overcome problems. Participants shared that most of them received support from their immediate and extended families. Considering individuals live in a collectivist culture in Turkey, family relations may have a protective role. Their partners, older children, and parents provided help to ease challenges. Teachers and school counselors were also reported as main sources of support, which was different to other studies conducted during the pandemic (Dong et al., 2020; Garbe et al., 2020). It might be interesting to sought answer to the same research question in an individualistic culture to better understand the importance of family and school support in different cultures. Additionally, according to the present study findings, parents sought creative ways to support their children. For instance, some participants reported contacting with other families to arrange regular outdoor meetings in playgrounds for children peer support. Thus, one can conclude that during ERT, parents worked extra hard to find ways to support their children as best as they can.

In the third and last research question, we tried to understand the changes in family relations during the ERT. All participants stated that their family relations were affected in both positive and negative ways. On one hand, participants reported feeling happy and grateful to be with their family members in a difficult time such as the pandemic. On the other hand, participants also reported difficulties such as running out of patience from time to time. Difficult times may bring difficult emotions and restrict coping mechanisms. However parents patience might reach their limits and they may need extra support in such circumstances. We did not ask our participants if they receive psychological support to keep the balance in their life and relationships. In future studies, this might be a further concept to investigate.

This study is not without limitation. First, recruiting participants from social media created a sampling bias that parents who did participate were probably much more technology-literate and technology-resourced than parents who were not on social media. Therefore, the findings should be considered with this limitation in mind. Additionally, we included parents with children at primary school level (1st, 2nd, 3rd and 4th graders). However, first graders may experience unique challenges as they learn how to read and write during ERT without any prior school experience. Future researchers may design an in-depth qualitative study with only first graders' parents to gain a greater understanding about their unique experiences.

The world is changing and technology is evolving consistently. We may not know when the next pandemic will occur, but we believe that taking advantage of remote teaching and distance education, when necessary, will be considered as an option. Therefore, parents need to be prepared for constant changes in life, and teachers and school counselors should consider what kind of support (digital, face to face, etc.) should be provided to parents in difficult times. We learned that parents may take new roles in a crisis, so schools should be aware of such tendency to ease this transition for parents. With this study, we hoped to provide a snapshot of Turkish parents personal experiences, problems, supports and family relations during the ERT in COVID-19. We also hope that being aware of potential problems, changes in family relations and support mechanisms may provide insights for other parents living in different countries, individuals who work with parents, and future researchers.

CONCLUSION AND RECOMMENDATIONS

Here are some implications for counselors and psychologist in the schools;

Parents developed a new routine to balance work and childcare, so they had to find their way in this difficult time and support their children as best as they could. So, it seems that there is a new gap for psychologists and counsellors in schools to provide guidance and assistance for parents on that new routine and ways to take care of their own wellbeing.

Children's limited socialization became a major problem, as socialization is vital for primary school children. As a consequence of limited socialization, screen time usage of children increased exceedingly. Therefore, psychologists and counsellors in schools should work with students and their parents to teach health ways to utilize technology. For instance, the Digital Citizenship Curriculum (<https://www.common sense.org/education/digital-citizenship>) which is a research based online curriculum can be introduced to families.

Despite the fact that it is insufficient, school counsellors and psychologists were the primary sources of mental health support for parents and children. So, they should enhance their ability to provide digital support by increasing their knowledge about digital mental health services.

Declaration of Conflicting Interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The authors received no financial support for the research, authorship, and/or publication of this article.

Statements of publication ethics

We hereby declare that the study has not unethical issues and that research and publication ethics have been observed carefully.

Researchers' contribution rate

The study was conducted and reported with equal collaboration of the researchers.

Ethics Committee Approval Information

Ethical approval was obtained from Uşak University Social and Humanities Ethical Board on 12/02/2021.

REFERENCES

- Aydın, C. H., 2011. *Açık ve uzaktan öğrenme: öğrenci adaylarının bakış açısı*. Pegem Akademi.
- Balls, P. (2009). Phenomenology in nursing research: methodology, interviewing and transcribing. *Nursing Times*, 105(32-33), 30-33.
- Bozkurt, A., & Sharma, R. C. (2020). Emergency remote teaching in a time of global crisis due to CoronaVirus pandemic. *Asian Journal of Distance Education*, 15(1), i-vi. DOI: 10.5281/zenodo.3778083
- Chamberlain, B. (2009). Phenomenology: A qualitative method. *Clinical Nurse Specialist*, 23(2), 52-53.
- Creswell, J. W., & Poth, C. N. (2016). *Qualitative inquiry and research design: Choosing among five approaches*. Sage publications.
- Denzin, N. K. (2017). *The research act: A theoretical introduction to sociological methods*. Transaction publishers.
- Dong, C., Cao, S., & Li, H. (2020). Young children's online learning during COVID-19 pandemic: Chinese parents' beliefs and attitudes. *Children and Youth Services Review*, Published online, 118, 105440. <https://doi.org/10.1016/j.childyouth.2020.105440>
- Garbe, A., Ogurlu, U., Logan, N., & Cook, P. (2020). Parents' experiences with remote education during COVID-19 school closures. *American Journal of Qualitative Research*, 4(3), 45–65. <https://doi.org/10.29333/ajqr/8471>.
- Golafshani, N. (2003). Understanding reliability and validity in qualitative research. *The Qualitative Report*, 8(4), 597-607. <https://doi.org/10.46743/2160-3715/2003.1870>
- Gürmen, M. S., & Kılıç, S. (2022). *Parenting in Turkey*. In Parenting Across Cultures: Childrearing, Motherhood and Fatherhood in Non-Western Cultures (pp. 243-257). Cham: Springer International Publishing.
- Hodges, C., Moore, S., Lockee, B., Trust, T., & Bond, A. (2020). The difference between emergency remote teaching and online learning. *Educause Review*, 27. <https://er.educause.edu/articles/2020/3/the-difference-between-emergency-remote-teaching-and-online-learning>
- Mengi, A. & Alpdoğan, Y. (2020). Covid-19 salgını sürecinde özel eğitim öğrencilerinin uzaktan eğitim süreçlerine ilişkin öğretmen görüşlerinin incelenmesi. *Milli Eğitim Dergisi*, 49(1), 413-437. DOI: 10.37669/milliegitim.776226
- Ministry of National Education (MoNE). (2020). *National education statistics: Formal education 2019/'20*. http://sgb.meb.gov.tr/meb_iys_dosyalar/2020_09/04144812_meb_istatistikleri_orgun_egitim_2019_2020.pdf
- Ministry of National Education (MoNE). (2021). *National Education Statistics Formal Education 2020-'21*. https://sgb.meb.gov.tr/meb_iys_dosyalar/2021_09/10141326_meb_istatistikleri_orgun_egitim_2020_2021.pdf
- Misirli, O., & Ergulek, F. (2021). Emergency remote teaching during the COVID-19 pandemic: Parents experiences and perspectives. *Education and Information Technologies*, 26, 6699–6718.
- Moustakas, C. (1994). *Phenomenological research methods*. SAGE Publications

- Munhall, P. L. (2007). Language and nursing research: The evolution. *Nursing research: A qualitative perspective*, 37-69.
- Parczewska, T. (2020). The quality of care for 5–9-year old children in the school environment in Poland measured with the use of the SACERS scale. *Education*, 48(5), 541-549.
- Patrick, S. W., Henkhaus, L. E., Zickafoose, J. S., Lovell, K., Halvorson, A., Loch, S., Letterie, M., & Davis, M. M. (2020). Well-being of parents and children during the COVID-19 pandemic: A national survey. *Pediatrics*, 146(4), e2020016824. <https://doi.org/10.1542/peds.2020-016824>
- Patton, M. Q. (2015). *Qualitative research and evaluation methods* (4th ed.). SAGE Publications.
- Pew Research Center. (2020, April 7) https://www.pewsocialtrends.org/wpcontent/uploads/sites/3/2020/04/PSDT_04.21.20_covidfinance_TOPLINE.pdf
- Plowman, L., Stevenson, O., Stephen, C., & McPake, J. (2012). Preschool children's learning with technology at home. *Computers & Education*, 59(1), 30-37. <http://dx.doi.org/10.1016/j.compedu.2011.1>
- Snelling, J., & D. Fingal. (2020). *10 strategies for online learning during a coronavirus outbreak*. International Society for Technology in Education. <https://www.iste.org/explore/10-strategies-online-learning-during-coronavirus-outbreak>.
- Trzcińska-Król, M. (2020). Students with special educational needs in distance learning during the COVID-19 pandemic—parents' opinions. *Interdyscyplinarne Konteksty Pedagogiki Specjalnej*, 29, 173-191. <https://doi.org/10.14746/ikps.2020.29.08>
- UNESCO. (2020, June 12). *COVID-19 Education Response*. <https://en.unesco.org/covid19/educationresponse/globalcoalition>
- UNESCO (2022). *Education: From disruption to recovery: COVID-19 impact on education*. <https://webarchive.unesco.org/web/20220629024039/https://en.unesco.org/covid19/educationresponse/>
- Van der Mescht, H. (2004). Phenomenology in education: A case study in educational leadership. *Indo-Pacific Journal of Phenomenology*, 4(1), 1-16.
- Weaver, J. L., & Swank, J. M. (2021). Parents' lived experiences with the COVID-19 pandemic. *The Family Journal*, 29(2), 136-142. <https://doi.org/10.1177/1066480720969194>
- Williamson, B., Eynon, R., & Potter, J. (2020). Pandemic politics, pedagogies and practices: Digital technologies and distance education during the coronavirus emergency. *Learning, Media and Technology*, 45(2), 107–114.
- Wu, M., Xu, W., Yao, Y., Zhang, L., Guo, L., Fan, J., & Chen, J. (2020). Mental health status of students' parents during COVID-19 pandemic and its influence factors. *General Psychiatry*, 33(4) 1-9. <http://dx.doi.org/10.1136/gpsych-2020-100250>



Factors Related to the Mathematics Achievement of Resilient and Nonresilient Students with Different Genders in Top Performing Asian Countries

Yüksek Başarılı Asya Ülkelerinde Farklı Cinsiyetlerden Dirençli ve Dirençsiz Öğrencilerin Matematik Başarılarıyla İlişkili Faktörler¹

Musa SADAK²

Keywords

1. Resilient and nonresilient students
2. Gender
3. Mathematics achievement
4. Multiple linear regression

Anahtar Kelimeler

1. Dirençli ve dirençsiz öğrenciler
2. Cinsiyet
3. Matematik başarıları
4. Çoklu doğrusal regresyon

Received/Başvuru Tarihi
26.05.2025

Accepted / Kabul Tarihi
10.10.2025

Abstract

Purpose: This study focused on the relationship between student-related factors and students' mathematics achievement in top performing five Asian countries, including, Singapore, Chinese Taipei, South Korea, Japan, and Hong Kong, in TIMSS 2019 at the eighth-grade level with the specific focus on resilient and nonresilient students with different genders.

Design/Methodology/Approach: The data consisted of the pooled sample of 19,652 students in total. Seven student-related factors indicated by previous researchers were tested across the resilient and nonresilient students in these Asian countries through linear multiple regression technique.

Findings: As a result, valuing of mathematics and liking mathematics were found to have significant positive relationship with resilient students' mathematics achievement while self-confidence and SES to be significantly related to only resilient male students' achievement. In addition, for the nonresilient students, valuing mathematics, SES, liking mathematics, and discipline climate of classroom were found to be significantly positively related to achievement for both genders.

Highlights: The study revealed that valuing and liking mathematics consistently contributed to higher mathematics achievement among both resilient and nonresilient students across genders. While self-confidence in mathematics and socioeconomic status (SES) were not significant for resilient female students, they emerged as important predictors for resilient male students as well as for both genders of nonresilient students. Moreover, classroom discipline climate positively influenced the achievement of nonresilient students but showed no effect on resilient students. In contrast, school belonging and experiences of bullying were not significantly related to mathematics achievement in the overall sample.

Öz

Çalışmanın amacı: Bu çalışma, öğrenciye ilişkin faktörler ile öğrencilerin matematik başarıları arasındaki ilişkiyi, TIMSS 2019 sekizinci sınıf düzeyinde Singapur, Çin Taipei, Güney Kore, Japonya ve Hong Kong gibi başarı düzeyi yüksek beş Asya ülkesinde, dirençli ve dirençsiz öğrenciler ile cinsiyet farklılıklarına odaklanarak incelemiştir.

Materyal ve Yöntem: Veriler, toplamda 19.652 öğrenciden oluşan havuz örneklemini kapsamaktadır. Önceki araştırmalarda belirtilen öğrenciye ilişkin yedi faktör, bu Asya ülkelerinde dirençli ve dirençsiz öğrenciler arasında çoklu doğrusal regresyon yöntemiyle test edilmiştir.

Bulgular: Sonuç olarak, matematiğe değer verme ve matematiği sevmenin, dirençli öğrencilerin matematik başarılarıyla anlamlı düzeyde pozitif ilişkili olduğu; öz güven ve sosyoekonomik statünün ise yalnızca dirençli erkek öğrencilerin başarılarıyla anlamlı düzeyde ilişkili olduğu bulunmuştur. Ayrıca, dirençsiz öğrenciler için, matematiğe değer verme, sosyoekonomik statü, matematiği sevmek ve sınıf disiplin ortamı hem kız hem erkek öğrencilerin başarılarıyla anlamlı ve pozitif yönde ilişkili bulunmuştur.

Önemli Vurgular: Çalışma, matematiğe değer verme ve matematiği sevmenin, hem dirençli hem de dirençsiz öğrencilerin farklı cinsiyetlerdeki başarılarını tutarlı bir biçimde artırdığını ortaya koymuştur. Matematikte öz güven ve sosyoekonomik statü (SES), dirençli kız öğrenciler için anlamlı bulunmazken, dirençli erkek öğrenciler ve dirençsiz öğrencilerin her iki cinsiyeti için önemli yordayıcılar olarak öne çıkmıştır. Ayrıca, sınıfın disiplin iklimi dirençsiz öğrencilerin başarılarını olumlu yönde etkilerken, dirençli öğrenciler üzerinde anlamlı bir etkisi bulunmamıştır. Buna karşın, okula aidiyet ve zorbalık deneyimleri genel örneklemede matematik başarıları ile anlamlı bir ilişki göstermemiştir.

¹ This study was presented as an oral presentation and published as an abstract paper at the 5th Educational Research Congress held on October 18–19, 2024

² **Corresponded Author**, Kastamonu University, Faculty of Education, Department of Educational Sciences, Kastamonu, Türkiye; <https://orcid.org/0000-0001-6036-1279>

INTRODUCTION

Numerous studies have been conducted to investigate factors associated with student achievement. While the factors are compiled around constructs as teacher-related, family-related, school-related, etc., researchers also brought student-related factors into attention, such as attitudes toward mathematics (e.g. Abazaoğlu, Yatağan, Yıldızhan, Arifoğlu, & Umurhan, 2015; Doğan & Barış, 2010; Uzun, Bütün, & Yiğit, 2010), valuing of mathematics (e.g. Atar, 2011; Doğan & Barış, 2010; Yayan & Berberoğlu, 2004), self-confidence in mathematics (e.g. Abazaoğlu et al., 2015; Akyüz, 2014; Atar, 2011; Doğan & Barış, 2010; Güven & Çabakçor, 2013; Yayan & Berberoğlu, 2004), school belonging (e.g. Akyüz & Pala, 2010; Engin-Demir, 2009), and home educational resources (e.g. Akyüz, 2006; 2014; Kılıç, Çene, & Demir, 2012; Özer & Anıl, 2011; Yayan & Berberoğlu, 2004). Among these factors, students' socio-economic status (SES) has also been getting attention of the researchers for a long time. Many studies have reported that students with higher SES are more likely able to achieve higher than their peers (Abazaoğlu et al., 2015; Akyüz, 2014; Akyüz & Pala, 2010; Alacacı & Erbaş, 2010; Demir, Kılıç, & Ünal, 2009; 2010; Dinçer & Kolasın, 2009; Dinçer & Uysal, 2010; Engin-Demir, 2009; Kılıç et al., 2012; Nilsen, Blömeke, Hansen, & Gustafsson, 2016; Özer & Anıl, 2011; Reardon, 2011; Xie & Ma, 2019). Even though students from families with lower SES are more likely to achieve poorly, many of these students are not vulnerable to their circumstances and still perform well in school, whom are called "academically resilient" since they overcome their disadvantageous situation (Agasisti, Avvisati, Borgonovi, & Longobardi, 2018; Agasisti & Longobardi, 2014; Avcı, 2022; Erberber, Stephens, Mamedova, Ferguson, & Kroeger, 2015; Martin & Marsh, 2006; Masten, 2014; OECD, 2018; Sandoval-Hernandez & Bialowolski, 2016). While the interestingness of these students seems quite an issue, identifying the reasons behind their unexpected success is still challenging for researchers (Erberber et al., 2015; Radisic & Petterson, 2020).

When examining the relevant literature in terms of revealing the characteristics that lead resilient students to be successful in mathematics at school, it was encountered with numerous studies, in which the factors indicated are mostly overlapped with the ones indicated in the studies mentioned above. The characteristics that are possibly linked to success of resilient students by researchers include their attitude toward mathematics (Frempong Visser, Feza, Winnaar, & Nuamah, 2016; Martin & Marsh, 2006; Sandoval-Hernandez & Bialowolski, 2016), valuing of mathematics (Erberber et al., 2015), self-confidence in mathematics (Martin and Marsh, 2009; OECD, 2012; Sandoval-Hernandez, & Cortes, 2012), relationships with their peers (Doll, Zucker, & Brehm, 2004; Lessard, Butler-Kisber, Fortin, & Marcotte, 2014), bullying they are faced with (Erberber et al., 2015), school climate in which they are instructed (Kyriakides, Creemers, Antoniou, & Demetriou, 2010; Martin, Foy, Mullis, & O'Dwyer, 2013; Maxwell, Reynolds, Lee, Subasic, & Bromhead, 2017; OECD, 2018), and disciplinary climate of their schools (Agasisti et al., 2018; Güzel & Berberoğlu, 2005; Ma, Jong, & Yuan, 2013; OECD, 2018; Shin, Lee, & Kim, 2009). Discipline issues are reported as an essential factor of resilience especially in Asian countries (Ma et al., 2013; Shin et al., 2009) where disciplined class environment is defined a class with free of disruption (Kyriakides & Creemers, 2008; Hopkins, 2005).

Research on resilience in terms of the gender differences revealed opposite results. While some of these indicated that female students are more resilient than male students (Allan, McKenna, & Dominey, 2014; DuMont, Widom, & Czaja, 2007; McLafferty, Mallet, & McCauley, 2012; Parker, Hogan, Easabrook, Oke, & Wood, 2006; Sun & Stewart, 2007), some indicated the opposite (Li, 2008; Martin & Marsh, 2006; 2008; Sander & Sanders, 2009). Although there are studies focused on the differences in resilience of different genders, there is a lack of studies focusing on revealing the factors behind the difference of resiliency of students from different genders, with the best understanding on a relevant literature review. Therefore, this study aimed to investigate how student-related factors are associated differently with mathematics achievement between resilient students with different genders. The research question shaping the form of the research is:

1. To what extent student-related factors indicated by previous researchers (SES, school belonging, bullying, liking mathematics, discipline climate of classroom, self-confidence in mathematics, and valuing of mathematics) are related with mathematics achievement of resilient and nonresilient students with different genders in top performing five Asian countries in TIMSS 2019?

METHOD

Sample

The sample of this study is consisted of eighth-grade students of top performing five Asian countries of mathematics part of TIMSS 2019 assessment, namely Singapore, Chinese Taipei, Korea Republic, Japan, and Hong Kong SAR. Thirty-nine countries and 7 jurisdictions were participated voluntarily in TIMSS 2019 assessment while the five Asian countries mentioned above are the top achievers by substantial margin (Mullis, Martin, Foy, Kelly, & Fishbein, 2020). These countries are the specific focus of this study due to the high standards and severe academic pressure that define these educational systems. The ways in which kids demonstrate resilience in such settings may be different from those in systems with mediocre or low performance. A greater comprehension of how students perform under demanding academic conditions is made possible by analyzing resilience in these circumstances. Therefore, pooled sample of these countries was examined for the groups of resilient female, resilient male, nonresilient female and nonresilient male students. When defining these groups, TIMSS 2019 "Intermediate International Benchmark of mathematics" was used as a cut score to determine successful students while "few resources" category of the variable "Home Educational Resources" was used as a cut score to determine advantaged students, as Erberber et al. (2015) described. As a result, students who perform above 475 points at mathematics assessment and are classified in or below few

resources category consist of resilient students in each individual country. Likewise, students who perform above 475 points at mathematics assessment and are classified above few resources category or who perform below 475 points at mathematics assessment and are classified in or below few resources category consist of nonresilient students. Sample sizes for each of these individualized groups are given below in Table 1.

Table 1. Sample sizes for resilient and nonresilient students by gender

	Resilient Students (Ach.>475 & HER<=few resources)	Nonresilient Students (Ach.>475 & HER>few resources) or (Ach.<475 & HER<=few resources)
Female	587	9202
Male	663	9200
Total	1250	18402

Data Sources

In TIMSS 2019 assessment, each student received an achievement booklet to assess their mathematical knowledge while they received a questionnaire to determine the factors behind their achievement (Mullis, 2017). The determination of the variables used to assess factors associated with student achievement was based on the relevant literature review. Table 2 shows the variables used in this study. While student achievement, which was the outcome variable, was assessed through a plausible value created by TIMSS, remaining variables, explanatory variables, were assessed through relevant variable created by TIMSS. In addition to the results for the singular questionnaire items posed to students, TIMSS developers also created some scaled variables that were derived from the combination of the singular items. In this study, these scaled variables used to assess student-related factors of their achievement. Table 2 shows these scaled variables along with the items used to derive them and the cut scores for the categories created to scale them. The data for this study were obtained online from IEA's website for the TIMSS 2019 International Data Sets.

Table 2. Student-related variables and correspondent items in TIMSS 2019

Variable	Relevant items in TIMSS 2019
SES (BSBGHER)	Number of books in the home: – 0-10, – 11-25, – 26-100, – 101-200, – >200.
	Number of home study supports: – None, – internet connection or own room, – both internet connection and own room.
	Highest level of education of either parent: – Finished some primary or lower secondary or did not go to the school, – Finished lower secondary, – Finished upper secondary, – Finished post-secondary education, – Finished university or higher.
School belonging (BSBGSSB)	What do you think about your school? Tell how much you agree with these statements (Agree a lot, agree a little, disagree a little, disagree a lot). – I like being in school. – I feel safe when I am at school. – I feel like I belong at this school. – Teachers at my school are fair to me. – I am proud to go to this school.
Bullying (BSBGSB)	During this school year, how often have other students from your school done any of the following things to you, including through texting or the Internet? – Made fun of me or called me names. – Left me out of their games or activities. – Spread lies about me. – Stole something from me. – Damaged something of mine on purpose. – Hit or hurt me (e.g., shoving, hitting, kicking) – Made me do things I didn't want to do. – Sent me nasty or hurtful messages online. – Shared nasty or hurtful messages about me online.

Variable	Relevant items in TIMSS 2019
	– Shared embarrassing photos of me online. – Threatened me.
Liking mathematics (BSBGSLM)	How much do you agree with these statements about learning mathematics (Agree a lot, agree a little, disagree a little, disagree a lot)? – I enjoy learning mathematics. – I wish I did not have to study mathematics. – Mathematics is boring. – I learn many interesting things in mathematics. – I like mathematics. – I like any schoolwork involving numbers. – I like to solve mathematics problems. – I look forward to mathematics lessons. – Mathematics is one of my favorite subjects.
Discipline climate of classroom (BSBGDML)	How often do these things happen in your mathematics lessons? – Students don't listen to what the teacher says. – There is disruptive noise. – It is too disorderly for students to work well. – My teacher has to wait a long time for students to quiet down. – Students interrupt the teacher. – My teacher has to keep telling us to follow the classroom rules.
Self-confidence in mathematics (BSBGSCM)	How much do you agree with these statements about mathematics (Agree a lot, agree a little, disagree a little, disagree a lot) (*reverse coded)? – I usually do well in mathematics. – Mathematics is more difficult for me than for many of my classmates*. – Mathematics is not one of my strengths*. – I learn things quickly in mathematics. – Mathematics makes me nervous*. – I am good at working out difficult mathematics problems. – My teacher tells me I am good at mathematics. – Mathematics is harder for me than any other subject*. – Mathematics makes me confused*.
Valuing of mathematics (BSBG SVM)	How much do you agree with these statements about mathematics (Agree a lot, agree a little, disagree a little, disagree a lot)? – I think learning mathematics will help me in my daily life. – I need mathematics to learn other school subjects. – I need to do well in mathematics to get into the university of my choice. – I need to do well in mathematics to get the job I want. – I would like a job that involves using mathematics. – It is important to learn about mathematics to get ahead in the world. – Learning mathematics will give me more job opportunities when I am an adult. – My parents think that it is important that I do well in mathematics. – It is important to do well in mathematics.

Data Analysis

In order to assess the extent to which student-related factors indicated by previous researchers are related with mathematics achievement of resilient and nonresilient students with different genders, linear multiple regression technique was used. While student achievement was used as an outcome variable, student-related factors indicated before were used as explanatory variables so that variability in students' mathematics achievement could be explained through student-related variables. This hypothesized model is given in Equation 1 below. After creating this hypothesized model, it was tested in the pooled sample of five Asian countries, which are the top achievers in TIMSS 2019. This multilevel model was estimated with a special package called *lme4* (Bates, 2022) in R statistical software (R Core Team, 2021).

$$Y_{ij} = \gamma_{00} + \sum_p \gamma_{p0} X_{pij} + r_{ij} \quad (1)$$

where

Y_{ij} is the mathematics achievement of a student i in the country j ;

γ_{00} is the intercept term;

γ_{p0} are the fixed regression coefficients representing the direction and strength of the relationship between student level variables (X_{pij}) and Y_{ij} ;

X_{pij} are student variables predicting achievement of student ij ; and

r_{ij} is a random “student effect” that is the deviation of student r_{ij} ’s mathematics achievement score from the estimated score.

Before testing the hypothesized model, it is very essential to check the assumptions of the linear multiple regression technique, which are linearity of the data, normality of residuals, homogeneity of residual variance (heteroscedasticity), and multicollinearity (Tabachnick & Fidell, 2013). Figure 1 and Table 3 show the result for the tests for each assumption for each individual model.

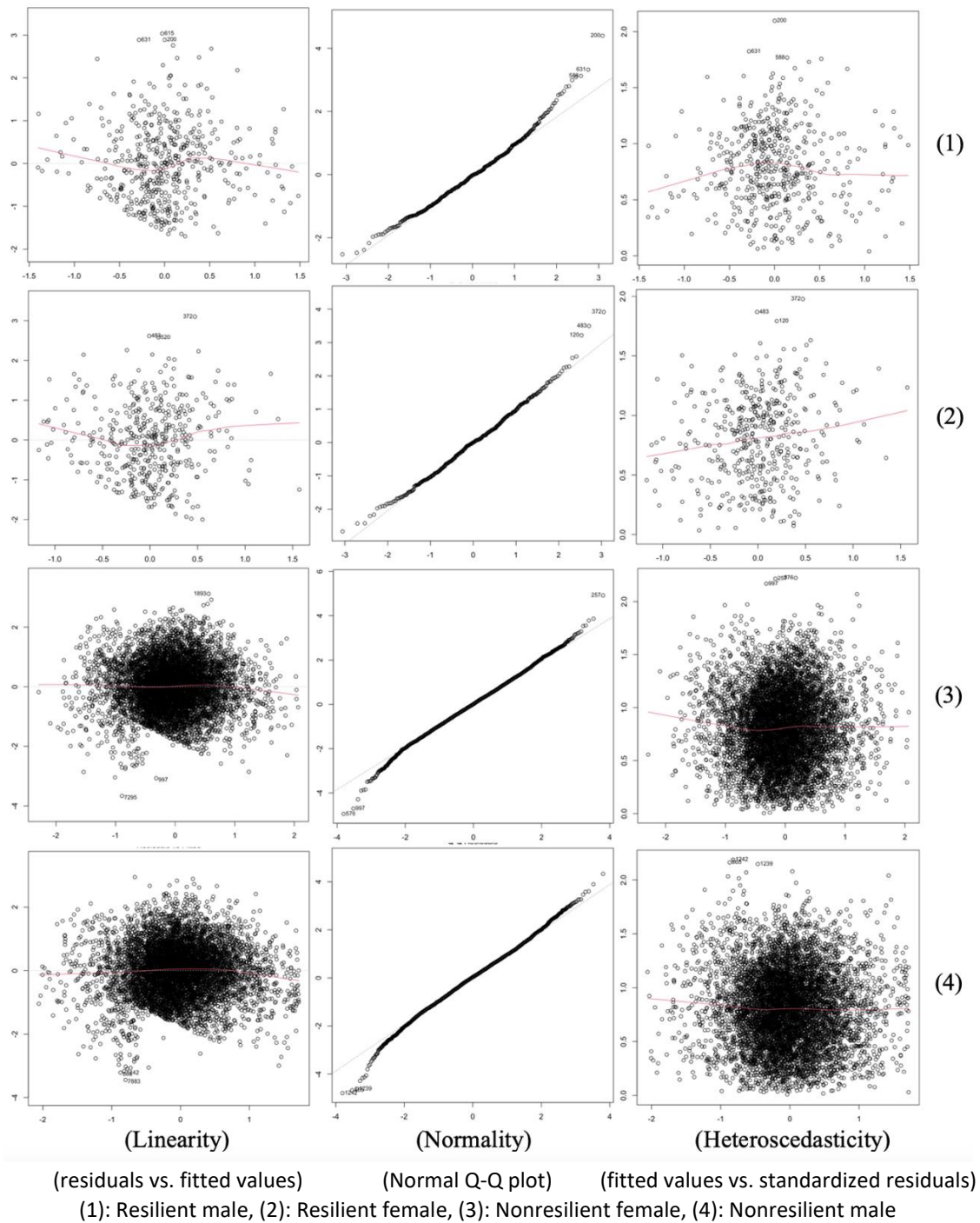


Figure 1. Plots for the assumptions of Linear Multiple Regression Models

As shown in Figure 1, plots for the residuals vs. fitted values show no fitted pattern across different models, which indicates that linearity assumption was met. On the other hand, normal Q-Q plots for the theoretical quantiles vs. standardized residuals approximately follow the straight line, which implies that the assumption of normality of residuals was met. In addition, plots for the fitted values vs. standardized residuals are spread approximately equally along the ranges of predictors, which suggests nonconstant variances in the residual errors, in other words heteroscedasticity assumption was met. Last but not the least, in addition to these three assumptions, the multicollinearity assumption was also checked. In order to check this assumption, bivariate correlations among the explanatory variables were examined (Table 3). According to Tabachnick and Fidell (2013), multicollinearity occurs when there is a bivariate correlation of greater than .90 between any two of the explanatory variables. As

shown in Table 3, the strongest correlation among the explanatory variables was observed to be .67, which indicates no issue in terms of multicollinearity.

Table 3. Bivariate correlations among explanatory variables

	BSBGHER	BSBGSSB	BSBGSB	BSBGSLM	BSBGDML	BSBGSCM
BSBGSSB	.05					
BSBGSB	-.05	.14				
BSBGSLM	-.00	.33	.03			
BSBGDML	-.00	.17	.30	.10		
BSBGSCM	.02	.10	-.00	.67	-.02	
BSBGSVM	.04	.37	-.05	.63	.03	.43

FINDINGS

In the linear multiple regression models, all regression coefficients were fixed and tested for resilient and nonresilient students with different genders in top performing five Asian countries in TIMSS 2019. Table 4 shows the results for each individual model. As shown, sample size for each sample decreased from the values given in Table 1 due to default missing data management of the statistical package by list-wise deleting missing values.

Table 4. Results from the Linear Multiple Regression Models with Mathematics Achievement as the Dependent Variable

	Resilient Male	Resilient Female	Nonresilient Male	Nonresilient Female
Variables	Estimate (S.E.)	Estimate (S.E.)	Estimate (S.E.)	Estimate (S.E.)
γ_{00} = Intercept	-.001 (.041)	.042 (.046)	-.020 (.010)	.024* (.010)
γ_{10} = SES	.079* (.040)	.074 (.046)	.315** (.010)	.297** (.010)
γ_{20} = School belonging	.004 (.044)	-.050 (.047)	-.006 (.011)	.007 (.011)
γ_{30} = Bullying	-.054 (.044)	.048 (.047)	-.018 (.011)	-.000 (.011)
γ_{40} = Liking math	.148* (.065)	.193** (.074)	.088** (.015)	.087** (.016)
γ_{50} = Disc. climate of class.	-.028 (.054)	.082 (.067)	.027* (.012)	.096** (.012)
γ_{60} = Self-confidence in math	.087* (.042)	-.020 (.046)	-.012 (.011)	.063 (.011)
γ_{70} = Valuing of math	.345** (.056)	.224** (.063)	.318** (.014)	.293** (.014)
N	498	436	7015	7054
Adjusted R-squared	.20	.16	.30	.30

* $p < .05$, ** $p < .01$. Standard errors were given in the parantheses. *Disc. climate of class.*: discipline climate of classroom

Results from the linear multiple regression models indicated that the extent to which covariates explain the total variance in students' mathematics achievement are seemingly low, ranging between 16% and 30%; however, it is reasonable to have this level of r-squared values when considering the fact that there were only seven student-related variables included in the models. As shown in Table 4, significance and magnitude of the relationships between explanatory variables and the outcome variable differs among different models conducted for resilient and nonresilient students with different genders. Since regression coefficients were estimated by the standardized measures, it is reasonable to compare them amongst each other in terms of the magnitude of their relationships with student achievement (Lorah, 2018).

As a result of the linear multiple regression models for the resilient students, it was observed that *valuing of mathematics* (γ_{70} -male = .345, $p < .01$; γ_{70} -female = .224, $p < .01$) and *liking mathematics* (γ_{40} -male = .148, $p < .05$; γ_{40} -female = .193, $p < .01$) had the first- and second-most strong significant positive relationship either with male or female resilient students' mathematics achievement. In addition, *self-confidence in mathematics* (γ_{60} -male = .087, $p < .05$) and *SES* (γ_{10} -male = .079, $p < .05$) were the other two significant positive predictors for students' mathematics achievement only for resilient male students, not for the resilient female. However, other student-related variables (*school belonging*, *bullying*, and *discipline climate of classroom*) were not found to have significant relationship neither with male nor female resilient students' mathematics achievement.

When examining the results of the linear multiple regression models for the nonresilient students, it was observed that *valuing of mathematics* (γ_{70} -male = .318, $p < .01$; γ_{70} -female = .293, $p < .01$), *SES* (γ_{10} -male = .315, $p < .01$; γ_{10} -female = .297, $p < .01$), *liking mathematics* (γ_{40} -male = .088, $p < .01$; γ_{40} -female = .087, $p < .01$) and *discipline climate of classroom* (γ_{50} -male = .027, $p < .05$; γ_{50} -female = .096, $p < .01$) were the ones having significant positive relationship with nonresilient male and female students' mathematics achievement although the order of the magnitude of the relationships of these variables with student achievement vary between male and female groups.

DISCUSSION, CONCLUSION AND RECOMMENDATIONS

According to some researchers, students with higher SES are more likely to perform higher than their peers (Abazaoğlu et al., 2015; Akyüz, 2014; Akyüz & Pala, 2010; Alacacı & Erbaş, 2010; Demir, Kılıç, & Ünal, 2009; 2010; Dinçer & Kolasin, 2009; Dinçer & Uysal, 2010; Engin-Demir, 2009; Kılıç et al., 2012; Nilsen, Blömeke, Hansen, & Gustafsson, 2016; Özer & Anil, 2011; Reardon, 2011; Xie & Ma, 2019). However, students with lower SES are not always vulnerable to their circumstances and achieve higher than their peers in some instances (Agasisti, Avvisati, Borgonovi, & Longobardi, 2018; Agasisti & Longobardi, 2014; Avcı, 2022; Erberber, Stephens, Mamedova, Ferguson, & Kroeger, 2015; Martin & Marsh, 2006; Masten, 2014; OECD, 2018; Sandoval-Hernandez & Bialowolski, 2016). Since identifying the reasons behind this unexpected situation gets attention (Erberber et al., 2015; Radisic & Petterson, 2020), in addition to the lack of research especially on gender differences, this study focused on revealing the differences between resilient and nonresilient students with different genders in terms of their mathematics achievement in top performing Asian countries in TIMSS 2019.

As a result of this study, *valuing of mathematics* and *liking mathematics* found to have significantly positive relationships with mathematics achievement of either resilient or nonresilient students with different genders, as previous researchers suggested (Erberber et al., 2015; Frempong Visser, Feza, Winnaar, & Nuamah, 2016; Martin & Marsh, 2006; Sandoval-Hernandez & Bialowolski, 2016). However, *self-confidence in mathematics* and *SES* were not found to have significant relationship with mathematics achievement of resilient female students as opposed to previous researchers' findings (Martin and Marsh, 2009; OECD, 2012; Sandoval-Hernandez, & Cortes, 2012) while these variables had significant positive relationship with resilient male students' achievement. This raises an important concern that how mathematics performance of female resilient students is not significantly related to their self-confidence and SES. It was also found as a result of this study that self-confidence was not also found to have significant relationship with mathematics achievement for nonresilient students. This could mean that female resilient students overcome their SES issues differently than their male counterparts while they overcome self-confidence issues differently than everybody else. This is also supported by previous research that female students are more resilient than the male (Allan, McKenna, & Dominey, 2014; DuMont, Widom, & Czaja, 2007; McLafferty, Mallet, & McCauley, 2012; Parker, Hogan, Easabrook, Oke, & Wood, 2006; Sun & Stewart, 2007). As a result, dealing with self-confidence and SES issues could be the key factors on becoming more resilient. However, there is a need for an in-depth analysis on these matters.

In addition, *discipline climate of classroom* was found to have a significant positive relationship with students' mathematics achievement for nonresilient students but not for resilient students. This is exactly opposite of previous researchers' statement indicating that discipline issues are reported as an essential factor of resiliency in Asian countries (Ma et al., 2013; Shin et al., 2009). It is very interesting that discipline issue is a significant positive indicator of nonresilient students. When disruptiveness of a classroom decreases by 1 standard deviation (SD), students' mathematics achievement increases by .027 SD ($p < .05$) for the nonresilient male and .096 ($p < .01$) for the nonresilient female. Since nonresilient students were defined as successful with higher SES or not successful with lower SES, it could be possible that students who are not successful and having lower SES could be the ones having discipline problems in the classroom. This might be the reason of how they did not overcome their disadvantageous situation as their resilient peers did. These findings indicated that it should not only be examined the reasons behind the success of resilient and nonresilient students but should also be examined the gender differences so that it could be revealed how different gender groups of resilient students overcome their disadvantageous situation.

On the other hand, *school belonging* and *bullying* were found not to have significant relationship with mathematics achievement for the entire sample, as opposed to previous researchers' findings (Erberber et al., 2015; Kyriakides, Creemers, Antoniou, & Demetriou, 2010; Martin, Foy, Mullis, & O'Dwyer, 2013; Maxwell, Reynolds, Lee, Subasic, & Bromhead, 2017; OECD, 2018). Therefore, as a result, it is highly suggested to future researchers that they should specifically focus on why

1. *self-confidence in mathematics* and *SES* for female resilient students,
2. *discipline climate of classroom* for resilient students, *self-confidence in mathematics* for nonresilient students, and
3. *school belonging*, and *bullying* for either resilient or nonresilient students

are not significant predictors of mathematics achievement as opposed to what previous research suggested. Conducting in-depth analyses on these matters could reveal the reality behind the performance of different genders of resilient and nonresilient students. Especially revealing the connection of the self-confidence in mathematics and SES to the process of how female resilient students become more resilient could potentially help students with lower SES, their teachers, and educational administrators to enhance these students' mathematical understanding.

LIMITATIONS OF THE STUDY

Since this study used a cross-sectional multi-country data, it was not very possible to make causal inferences on the relationship between student-related variables and students' mathematics achievement. Therefore, the findings of this study would be beneficial to the researchers who are able to investigate these issues in the real classroom environment through in-depth analyses. Results from multi-country data would help them to see a bigger picture of the issues when conducting a smaller-scale study. In addition, responses of students to the questionnaire items may vary among different countries, which was defined as *measurement invariance* (Steenkamp & Baumgartner, 1998).

Declaration of Conflicting Interests

The author declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author received no financial support for the research, authorship, and/or publication of this article.

Statements of publication ethics

I hereby declare that the study has not unethical issues and that research and publication ethics have been observed carefully.

Ethics Committee Approval Information

Since the study involves no human subjects, whose identities can be identified, it was exempted from requiring ethical approval. It was used encrypted data (TIMSS 2019) from IEA's website.

REFERENCES

- Abazaoglu, I., Yatağan, M., Yıldızhan, Y., Arifoğlu, A., & Umurhan, H. (2015). Öğrencilerin Matematik başarısının uluslararası fen ve matematik eğilimleri araştırması sonuçlarına göre değerlendirilmesi [Examination of the students' mathematics achievement by the trends in international mathematics and science study]. *Electronic Turkish Studies*, 10(7), 33-49.
- Agasisti, T., & Longobardi, S. (2014). Inequality in education: Can Italian disadvantaged students close the gap? *Journal of Behavioral and Experimental Economics*, 52, 8-20.
- Agasisti, T., Avvisati, F., Borgonovi, F., & Longobardi, S. (2018). *Academic resilience: What schools and countries do to help disadvantaged students succeed in PISA*. OECD education working papers, 167. OECD Publishing.
- Akyüz, G., & Pala, N. M. (2010). The effect of student and class characteristics on mathematics literacy and problem solving in PISA 2003. *Elementary Education Online*, 9(2), 668-678.
- Akyüz, G. (2006). Investigation of the effect of teacher and class characteristics on mathematics achievement in Turkey and European Union countries. *Elementary Education Online*, 5(2), 75-86.
- Akyüz, G. (2014). The effects of student and school factors on mathematics achievement in TIMSS 2011. *Eğitim ve Bilim*, 39(172), 150-162.
- Alacacı, C., & Erbaş, A. K. (2010). Unpacking the inequality among Turkish schools: Findings from PISA 2006. *International Journal of Educational Development*, 30(2), 182-192.
- Allan, J. F., McKenna, J., & Dominey, S. (2014). Degrees of resilience: Profiling psychological resilience and prospective academic achievement in university inductees. *British Journal of Guidance & Counselling*, 42(1), 9-25.
- Atar, B. (2011). Application of descriptive and explanatory item response models to TIMSS 2007 Turkey mathematics data. *Eğitim ve Bilim*, 36(159), 255-269.
- Avcı, S. (2022). Investigation of the individual characteristics that predict academic resilience. *International Journal of Contemporary Educational Research*, 9(3), 543-556.
- Bates, D. (2024). *Linear mixed-effects models using 'Eigen' and S4*. R package (version 1.1-35.4). <https://cran.r-project.org/web/packages/lme4/lme4.pdf>. Accessed 11 June 2024.
- Demir, I., Kılıç, S., & Deprem, O. (2009). Factors affecting Turkish students' achievement in mathematics. *US-China Education Review*, 6(6), 47-53.
- Demir, I., Kılıç, S., & Ünal, H. (2010). Effects of students' and schools' characteristics on mathematics achievement: Findings from PISA 2006. *Procedia Social and Behavioral Sciences*, 2(2010), 3099-3103.
- Dinçer, M. A., & Kolasın, G. U. (2009). *Türkiye'de öğrenci başarısında eşitsizliğin belirleyicileri* [The determination of inequalities in student success in Turkey]. Istanbul, Turkey: Sabancı Üniversitesi Eğitim Girişimi Reformu.
- Dincer, M. A., & Uysal, G. (2010). The determinants of student achievement in Turkey. *International Journal of Educational Development*, 30(6), 592-598.
- Doğan, N., & Barış, F. (2010). Tutum, değer ve özyeterlik değişkenlerinin TIMSS-1999 ve TIMSS-2007 sınavlarında öğrencilerin matematik başarılarını yordama düzeyleri [The prediction of students' attitudes, values, and self-efficacies variables on their mathematics achievement on TIMSS-1999 and TIMSS-2007]. *Eğitimde ve Psikolojide Ölçme ve Değerlendirme Dergisi*, 1(1), 44-50.
- Doll, B., Zucker, S., & Brehm, K. (2004). *Resilient classrooms: Creating healthy environments for learning*. Guilford Publications.
- DuMont, K. A., Widom, C.S., & Czaja, S.J. (2007). Predictors of resilience in abused and neglected children grown-up: The role of individual and neighborhood characteristics. *Child Abuse and Neglect*, 31, 255-274.

- Engin-Demir, C. (2009). Factors influencing the academic achievement of the Turkish urban poor. *International Journal of Educational Development*, 29(2009), 17-29.
- Erberber, E., Stephens, M., Mamedova, S., Ferguson, S., & Kroeger, T. (2015). Socioeconomically disadvantaged students who are academically successful: Examining academic resilience cross-nationally. *IEA's Policy Brief Series*, 5. Amsterdam
- Frempong, G., Visser, M., Feza, N., Winnaar, L., & Nuamah, S. (2016). Resilient learners in schools serving poor communities. *Electronic Journal of Research in Educational Psychology*, 14(2), 352-367.
- Güven, B., & Çobakçor, B. O. (2013). Factors influencing mathematical problem-solving achievement of seventh grade Turkish students. *Learning and Individual Differences*, 23(2013), 131-137.
- Güzel, Ç. I., & Berberoğlu, G. (2005). An analysis of the Programme for International Student Assessment 2000 (PISA 2000) mathematical literacy data for Brazilian, Japanese, and Norwegian students. *Studies in Educational Evaluation*, 31, 283-314.
- Hopkins, D. (2005), *The practice and theory of school improvement: International handbook of educational change*, (Vol. 4), Springer Science and Business Media.
- Kılıç, S., Çene, E., & Demir, I. (2012). Comparison of learning strategies for mathematics achievement in Turkey with eight countries. *Educational Sciences*, 12(4), 2594-2598.
- Kyriakides, L., & Creemers, B. P. (2008). Using a multidimensional approach to measure the impact of classroom-level factors upon student achievement: A study testing the validity of the dynamic model. *School Effectiveness and School Improvement*, 19(2), 183-205.
- Kyriakides, L., Creemers, B., Antoniou, P., & Demetriou, D. (2010). A synthesis of studies searching for school factors: Implications for theory and research. *British Educational Research Journal*, 36(5), 807-830.
- Lessard, A., Butler-Kisber, L., Fortin, L., & Marcotte, D. (2014). Analysing the discourse of dropouts and resilient students. *The Journal of Educational Research*, 107(2), 103-110.
- Li, M.H. (2008). Helping college students cope: Identifying predictors of active coping in different stressful situations. *Journal of Psychiatry, Psychology and Mental Health*, 2(1), 1-15.
- Lorah, J. A. (2018). Effect size measures for multilevel models: Definition, interpretation, and TIMSS example. *Large-Scale Assessments in Education*, 6(1), 8.
- Ma, X., Jong, C., & Yuan, J. (2013). Exploring reasons for the East Asian success in PISA. In H. D. Meyer & A. Benavot (Eds.), *PISA, power, and policy: The emergence of global educational governance*, Oxford: Symposium Books.
- Martin, A. J., & Marsh, H. W. (2006). Academic resilience and its psychological and educational correlates: A construct validity approach. *Psychology in the Schools*, 43(3), 267-281.
- Martin, A. J., & Marsh, H. W. (2008). Academic buoyancy: Towards an understanding of students' everyday academic resilience. *Journal of School Psychology*, 46, 53-83.
- Martin, A.J. & Marsh, H. W. (2009). Academic resilience and academic buoyancy: Multidimensional and hierarchical conceptual framing of causes, correlates and cognate constructs. *Oxford Review of Education*, 35(3), 353-370.
- Martin, M.O., Foy, P., Mullis, I.V.S., & O'Dwyer, L.M. (2013). Effective schools in reading, mathematics, and science at the fourth grade. In M.O. Martin & I.V.S. Mullis (Eds.), *TIMSS and PIRLS 2011: Relationships among reading, mathematics, and science achievement at the fourth grade—Implications for early learning*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.
- Masten, A. S. (2014). Global perspectives on resilience in children and youth. *Child Development*, 85, 6-20.
- Maxwell, S., Reynolds, K., Lee, E., Subasic, E., & Bromhead, D. (2017). The impact of school climate and school identification on academic achievement: Multilevel modeling with student and teacher data. *Frontiers in Psychology*, 8, 1-21.
- McLafferty M., Mallet J., & McCauley V. (2012). Coping at university: The role of resilience, emotional intelligence, age and gender. *Journal of Quantitative Psychological Research*, 1, 1-6.
- Mullis, I. V. S. (2017). Introduction. In Mullis, I. V. S., Martin, M. O. (Eds.). *TIMSS 2019 assessment frameworks (pp. 1-10)*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center at Boston College.
- Mullis, I. V. S., Martin, M. O., Foy, P., Kelly, D. L., & Fishbein, B. (2020). *TIMSS 2019 International Results in Mathematics and Science*. Retrieved from Boston College, TIMSS & PIRLS International Study Center. Website: <https://timssandpirls.bc.edu/timss2019/international-results/>
- Nilsen, T., Blömeke, S., Hansen, K. Y., & Gustafsson, J. E. (2016). *Can schools contribute to enhance equity? The answer may depend on a country's developmental level*. IEA Policy Brief.
- OECD. (2012). *PISA 2012 Results: Excellence through Equity (Volume II): Giving Every Student the Chance to Succeed*. Paris: OECD Publishing.
- OECD. (2018). *Equity in education: Breaking down barriers to social mobility*. OECD Publishing.

- Özer, Y., & Anıl, D. (2011). Examining the factors affecting students' science and mathematics achievement with structural equation modeling. *Hacettepe Üniversitesi Eğitim Fakültesi Dergisi*, 41(2011), 313,324.
- Parker, J. D. A., Hogan, M. J., Eastabrook, J. M., Oke, A., & Wood, L. M. (2006). Emotional intelligence and student retention: Predicting the successful transition from high school to university. *Personality and Individual Differences*, 41, 1329-1336.
- R Core Team (2021). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL: <https://www.R-project.org/>.
- Radisic, J., & Pettersen, A. (2020). Resilient and nonresilient students in Sweden and Norway—Investigating the interplay between their self-beliefs and the school environment. In T. S. Frones, A. Pettersen, J. Radisic, N. Buchholtz (Eds.), *Equity, equality and diversity in the Nordic model of education*, Cham, Switzerland: Springer.
- Reardon, S. F. (2011). The widening academic achievement gap between the rich and the poor: New evidence and possible explanations. In R. Murnane & G. Duncan (Eds.), *Whither opportunity? Rising inequality in schools, and children's life changes*. Russel Sage Foundation Press.
- Sander, P., & Sanders, L. (2009). Measuring academic behavioural confidence: The ABC scale revisited. *Studies in Higher Education*, 34(1), 19–35.
- Sandoval-Hernandez, A., & Cortes, D. (2012). *Factors and conditions that promote academic resilience: A cross-country perspective*. Paper presented at the annual conference of the Comparative and International Education Society, Puerto Rico.
- Sandoval-Hernandez, A., & Bialowolski, P. (2016). Factors and conditions promoting academic resilience: A TIMSS-based analysis of five Asian education systems. *Asia Pacific Education Review*, 17, 511–520.
- Shin, J., Lee, H., & Kim, Y. (2009). Student and school factors affecting mathematics achievement: International comparisons between Korea, Japan, and the USA. *School Psychology International*, 30, 520–537.
- Steenkamp, J. B. E., & Baumgartner, H. (1998). Assessing measurement invariance in cross national consumer research. *Journal of Consumer Research* 25(1), 78–90.
- Sun, J., & Stewart, D. (2007). Age and gender effects on resilience in children and adolescents. *International Journal of Mental Health Promotion*, 9(4), 16–25.
- Tabachnick, B. G., & Fidell, L. S. (2013). *Using multivariate statistics (sixth edition)*. New Jersey: Pearson.
- Uzun, S., Bütüner, S. O., & Yiğit, N. (2010). A comparison of the results of TIMSS 1999-2007: The most successful five countries-Turkey sample. *Elementary Education Online*, 9(3), 1174-1188.
- Xie, C., & Ma, Y. (2019). The mediating role of cultural capital in the relationship between socio-economic status and student achievement in 14 economies. *British Educational Research Journal*, 45, 838–855.
- Yayan, B., & Berberoğlu, G. (2004). A re-analysis of the TIMSS 1999 mathematics assessment data of the Turkish students. *Studies in Educational Evaluation*, 30(2004), 87-104.



| Research Article / Araştırma Makalesi |

Academics' Perceptions of "Artificial Intelligence": A Metaphorical Review

Akademisyenlerin "Yapay Zekâ"ya İlişkin Algıları: Metaforik Bir İnceleme

Meryem Dilek¹, Ümit Dilekçi²

Keywords

Artificial intelligence
Academician
Metaphor

Anahtar Kelimeler

Yapay zekâ
Akademisyen
Metafor

Received/Başvuru Tarihi
04.01.2025

Accepted / Kabul Tarihi
10.10.2025

Abstract

Purpose: This study unveils academics' mental images regarding the concept of artificial intelligence (AI) through metaphors.

Methodology: The study group included 93 academics working at various universities in Türkiye and reached through convenient sampling. The data was collected through a semi-structured form. Content and descriptive analyses were adopted to analyse the data. The findings suggested that participants used 67 distinct metaphors and nine recurring metaphors to describe AI. Based on the justifications for the metaphors, the researchers identified categories. Then, the metaphors were ranked by their frequency and percentages.

Findings: The findings showed that metaphors fell within three categories which were positive, negative and ambivalence. In the positive category "Humanized Intelligence, Efficiency and Support, Learning and Knowledge Transfer, Innovation and Creativity"; in the negative category "Security Concerns, Loss of Control, Misleading, Ethical Concerns and Inequality" and ambivalence category "Power and Danger", Innovation and Uncertainty", "Efficiency and Risk" were the emerging themes. The most metaphors were used under the theme of "Efficiency and Support".

Highlights: This study reveals that academics generally perceive AI positively as a tool accelerating work processes and increasing efficiency. However, AI has negative aspects, such as security concerns, loss of control, ethical issues, and inequality. The study reveals that academics tend to make a balanced assessment of AI, carefully weighing its potential benefits against its associated risks. This nuanced perspective offers valuable insights that can inform the development of future policies and best practices regarding AI integration.

Öz

Çalışmanın amacı: Bu çalışmanın amacı, akademisyenlerin yapay zekâ (YZ) kavramına ilişkin sahip oldukları zihinsel imgeleri metaforlar aracılığı ile ortaya koymaktır.

Yöntem: Türkiye'nin çeşitli üniversitelerinde görev yapan ve kolay ulaşılabilir örnekleme ile belirlen 93 akademisyen araştırmanın katılımcı grubunu oluşturmaktadır. Araştırmada mecazlar yoluyla veri toplama deseni tercih edilmiştir. Araştırmacılar tarafından hazırlanmış olan yarı yapılandırılmış form kullanılarak toplanan verilerin analizinde içerik ve betimsel analiz yöntemleri uygulanmıştır. Yapılan analizler sonucunda katılımcıların ürettikleri metaforlardan 67'sinin benzersiz olduğu, 9'unun ise birden fazla kez tekrarlandığı belirlenmiştir. Katılımcıların metaforları oluştururken kullandıkları gerekçeler doğrultusunda YZ kavramına ilişkin kavramsal kategoriler belirlenmiş, daha sonra metaforlar sıklık ve yüzdelik derecesine göre sıralanmıştır.

Bulgular: Yapılan analizler neticesinde metaforların olumlu, olumsuz, hem olumlu hem de olumsuz olmak üzere üç ana kategori altında toplandığı görülmüştür. Olumlu kategorisinde 4 tema (İnsanlaştırılmış Zekâ, Verimlilik ve Destek, Öğrenme ve Bilgi Aktarımı, Yenilik ve Yaratıcılık) olumsuz kategorisinde 4 tema (Güvenlik Sorunu, Kontrol Kaybı, Yanıltıcılık, Etik Sorunu ve Eşitsizlik) ve hem olumlu hem de olumsuz kategorisinde ise 3 tema (Güç ve Tehlike, Yenilik ve Belirsizlik, Verimlilik ve Risk) ortaya çıktığı görülmüştür. YZ'ye ilişkin görüşlerde olumlu kategorisinde en çok "Verimlilik ve Destek", olumsuz kategorisinde "Kontrol Kaybı", hem olumlu hem de olumsuz kategorisinde ise "Yenilik ve Belirsizlik" öne çıkan temalar olmuştur. Bütün temalar (n=11) birlikte değerlendirildiğinde ise YZ'ye yönelik algıların en çok "Verimlilik ve Destek" teması altında yoğunlaştığı görülmüştür.

Önemli Vurgular: Bu çalışma akademisyenlerin YZ'yi genellikle iş sürecini hızlandırıcı ve verimliliği artıran bir araç olarak olumlu bir şekilde algıladıklarını ancak YZ'nin güvenlik, kontrol kaybı, etik sorunlar ve eşitsizlik gibi olumsuz yönleri olduğunu da ortaya koymaktadır. Araştırmada akademisyenler tarafından YZ'nin faydaları ve potansiyel riskleri açısından görece dengeli bir değerlendirme yapıldığı görülmüştür. Bunun da gelecekteki politika ve uygulamalar için önemli bir perspektif sunacağı düşünülmektedir.

¹ Res. Assist., Batman University, Batman, TÜRKİYE; <https://orcid.org/0000-0002-5623-9331>

² Corresponded Author Assoc. Prof. Dr., Batman University, Batman, TÜRKİYE; <https://orcid.org/0000-0002-6205-1247>

INTRODUCTION

Technology has been a key driver of change, profoundly influencing both social structures and individual lives (Chui, 2016; Schwab, 2017). In recent years, societies have undergone a significant transformation fuelled by the integration of modern technologies (Brynjolfsson & McAfee, 2014; Castells, 2011; West, 2018). At the heart of this transformation lies Artificial Intelligence (AI), a ground breaking technology enabling machines to think and learn in ways that resemble human cognition (Mitchell, 1997; Russell & Norvig, 2020).

"AI is defined as the ability of a system to interpret external data correctly, learn from this data, and apply what it has learned to achieve specific goals and tasks through flexible adaptation" (Haenlein & Kaplan, 2019). According to Kaplan and Haenlein (2020), defining AI is inherently challenging for three primary reasons. First, there is the difficulty of defining human intelligence itself and the added complexity of applying this already ambiguous concept to machines. Second, the so-called AI effect—where tasks performed by machines, initially deemed intelligent, become commonplace and are no longer perceived as signs of intelligence—further complicates the definition. Finally, confusion surrounding the various stages and types of AI adds another layer of difficulty to providing a clear and consistent definition.

The contributions of leading scientist in the 20th century paved the way for a significant turn in research on AI, a complex and rapidly evolving field with various definitions reflecting its multifaceted nature. British mathematician and computer scientist Alan Turing is commonly regarded as having laid the groundwork for AI in its modern sense. In his seminal paper in 1950, *"Computing Machinery and Intelligence"*, Turing alleged that computers could acquire the ability to think like humans. He introduced the now-famous *"Turing Test,"* designed to assess whether a machine can display human-like intelligence. In this test, an interrogator communicates with a human and a computer via written exchanges, attempting to discern which is which. If the machine's responses are indistinguishable from the human's, Turing proposed that it could be considered to demonstrate *"intelligence"* (Copeland, 2024).

The concept of *"Artificial Intelligence"* (AI) was first formally introduced in 1956 by John McCarthy and his collaborators at the *Dartmouth Summer Research Project on Artificial Intelligence (DSRPAI)*, held at Dartmouth College in New Hampshire. This conference sought to bring together experts from various fields to explore a new domain dedicated to designing machines capable of emulating human intelligence (Kaplan & Haenlein, 2020). Between 1952 and 1969, International Business Machines (IBM) pioneered creating one of the earliest AI systems, a program designed to play chess. Around the same time, innovative projects like Arthur Samuel's checkers-playing software and Allen Newell and Herbert A. Simon's Logic Theorist considered the first actual AI program, laid foundational groundwork for the field. By the 1980s, AI had progressed into an industry, with artificial neural networks regaining attention in 1986, marking a significant shift in research and application (Ateş, 2024). In 1997, a critical breakthrough occurred when IBM's Deep Blue outperformed world chess champion Garry Kasparov, signalling the tangible capabilities of AI technologies (Schultz & Ellen-Schultz, 2007). Another turning point came in 2011 with Apple's release of Siri, a virtual assistant that brought advances in language comprehension and interaction to mainstream technology (Kırık & Özkoçak, 2023). Today, corporations such as OpenAI, Microsoft, Apple, Google, and Amazon are shaping the future of AI with tools and services woven into the fabric of modern life. These contributions redefine technological boundaries and accelerate the trajectory of AI's evolution across diverse sectors.

AI is increasingly integrated into the fabric of daily life, transforming sectors with its capacity to enhance productivity, facilitate decision-making, and contribute to teaching and learning. In education, AI is widely employed to streamline school and classroom management processes (Chiu et al., 2023; Görgülü et al., 2025; Holmes et al., 2019; Nartgün & Kennedy, 2024a; Polat, 2024; Sipahioğlu, 2023). In healthcare, it plays a critical role in diagnosing diseases and developing treatment plans (Par, 2024; Yenikaya & Oktaysoy, 2023; Yiğit et al., 2023). The finance sector utilizes AI for detecting fraud and crafting investment strategies (Özdemir, 2023; Yıldız, 2022), while the transportation industry applies it in the development and operation of autonomous vehicles (Ece, 2024; Şen & Gür, 2024). AI enhances user engagement in the media sector by delivering personalized content recommendations (Ateşgöz, 2023). Retail improves efficiency through optimized inventory management and responsive customer service systems (Güven & Ayvaz Güven, 2023). The judicial system, too, benefits from AI's capacity for analysing documents and assisting in case preparation (Efe, 2022).

Although AI brings numerous advantages, it also introduces complicated social and personal challenges tied to its adverse impacts. These include threats to privacy and confidentiality, disruptions in employment, widening social inequalities, and risks of biased decision-making (Abrams, 2024; Farahani & Ghasemi, 2024; Jobin et al., 2019; OECD, 2024). Privacy concerns emerge mainly from handling extensive datasets, where personal information becomes susceptible to breaches or cyber-attacks, compromising individuals' confidentiality (Akbulak, 2022; Eryılmaz, 2023; Mijwil et al., 2022). In the workplace, the automation facilitated by AI puts routine jobs at significant risk, with potential cumulative effect on economic stability and social cohesion (Sheikhi, 2022; WEF, 2020). Beyond employment, disparities in access to AI technology threaten to deepen existing income divides, especially between wealthier and developing nations, exacerbating global inequalities (Etiş, 2023; Joyce & Cruz, 2024). AI systems can magnify ethical dilemmas in sensitive areas like justice, healthcare, and recruitment. Often, these issues stem from reliance on historical data that embeds societal biases or developers unintentionally transferring their prejudices into the system's design (Barocas et al., 2019; Li et al., 2019; Zlateva et al., 2024).

Many sectors can experience the positive and negative aspects of AI simultaneously. For instance, in education, while AI provides benefits such as personalized learning for students and reduced workload for teachers, it also accommodates risks such as data security concerns, limitations of critical thinking skills, and reduced student-teacher interaction (Chiu et al., 2023; Holmes et al., 2019; Nartgün & Kennedy, 2024a). In this respect, it can be said that AI entails a transformation process that needs to be meaningfully integrated into its areas of use.

As in many fields, the academia is also going through a transformation process in the context of AI. For this transformation to be meaningful, it is considered essential to incorporate AI into academic work appropriately. AI applications in higher education have the potential to significantly improve learning and presentation experiences (Crompton & Song, 2021). Tasks performed by AI led academicians to question previous research and led to changes in research areas (Altıntop, 2023). For example, tasks such as article segmentation, analysis, and data processing can be done more quickly through AI (Mijwil et al., 2023). While this was initially considered a positive development, concerns such as reducing the human factor, replacing humans (Crompton & Burke, 2023), and plagiarism (Nartgün & Kennedy, 2024b) led to discussions.

Noy and Zhang (2023) conducted a study on individuals with a bachelor degree and investigated the effects of generative AI tools on professional writing tasks. The findings suggested that users of ChatGPT, a generative AI tool, were more effective, efficient, and motivated. In particular, the benefits it provided for low-skilled participants pointed to the potential of AI to close productivity gaps, while it was observed that it improved the work of high-skilled participants and reduced their task time. Another study conducted with academicians in Türkiye suggested that ChatGPT was regarded as a support tool by the participants (Livberber & Ayvaz, 2023). In particular, ChatGPT saves academic time, but plagiarism, and disinformation is prominent among academicians' ethical concerns. Such ethical concerns deepen the discussions on how AI can be used in a more proper way in academia. According to another study by Dolunay and Temel (2024), using AI technologies in academia facilitates article structuring, data analysis and translation. However, it also brings concerns such as producing inaccurate information, ethical problems and reducing the value of human authorship. Thus, the effects of AI resulted in many discussions from both positive and negative perspectives. These discussions offer clues about the further use of AI and how it will find a place in academia. Based on these discussions, it is significant to understand the perceptions of AI among academicians.

Purpose and Significance of the Study

AI became a technology widely discussed in education in recent years (Pedro et al., 2019; Sipahioğlu, 2022). Thus, it is significant to understand the perceptions of AI and the patterns underlying these perceptions. Previous research focused on technical aspects of AI (Niskanen et al., 2023), application areas (Keleş, 2022; Zawacki-Richter et al., 2019), and ethics (Huang et al., 2023). In addition, although there are studies on the relationship between AI and academia in the literature (Mijwil et al., 2023; Sariyasa & Monika, 2023; Thunström, 2022), to our best knowledge, there is no study addressing the metaphorical perceptions of academicians regarding AI. The primary purpose of this research is to explain academicians' views on AI through metaphors. In this respect, it is anticipated that the study will hold significance in making inferences regarding the understanding of perceptions regarding AI in academia.

Metaphors are defined as conceptual tools that shape our perception, thought processes and behavior, as well as decorative linguistic expressions. Metaphors reflect how individuals interpret and articulate complex concepts, providing valuable insights into their cognitive structures (Schmitt, 2005). How academicians perceive AI is critical in determining how it can be included in education and instructional practices. Accordingly, this research is expected to constitute an important reference point in the development and implementation of strategies regarding the role of AI in the academia. To this end, the study sought answers to the following questions:

1. *Which metaphors do academicians use to describe AI?*
2. *Under which themes do these metaphors fall?*

METHOD

Research Design

This study explores academicians' perceptions of AI through metaphors, a powerful tool for qualitative data collection in educational research (Aubusson, 2002; Midgley et al., 2014). Metaphors are particularly effective in uncovering hidden dimensions of participants' experiences and worldviews, making them an essential method in qualitative research (Schmitt, 2005). By drawing connections between new and prior experiences, metaphors provide a way to understand abstract concepts such as AI through analogy (Lakoff & Johnson, 1981). This study used metaphors to define and explain how academicians conceptualize AI. The analysis of these metaphors offers insights into the meanings academicians attribute to AI, following a qualitative research design (Yıldırım & Şimşek, 2016).

Participants

The study included 93 academicians working across various disciplines at universities in Türkiye. Participants were selected using convenience sampling, a method often employed when other sampling techniques are impractical. Convenience sampling offers speed and practicality, as it allows researchers to select accessible and nearby participants (Yıldırım & Şimşek, 2016). Of the

participants 57 were male and 36 were female. Regarding their discipline, the highest representation came from the Social, Human, and Administrative Sciences ($f=25$), followed by Educational Sciences ($f=20$). Additionally, 15 participants were from Health Sciences, and 8 were from Engineering. The remaining 25 participants were distributed across other disciplines, which ensured a diverse representation. This distribution is thought to provide the study with a broad and varied perspective and enrich the analysis of academicians' conceptualizations of AI.

Data Collection Tool and Data Collection

Before the preparation of the data collection tool, existing studies on AI using metaphors were reviewed (Aktaş, 2021; Demirtaş & Türksoy, 2023; Saçan et al., 2022). Then, the data collection tool was prepared by considering the common expressions in these studies. The data collection tool included two parts: The first part consists of questions aimed at identifying the demographics of academicians, and the second part consists of the expression "*Artificial intelligence is like..... because.....*," which aims to reveal the perceptions of academicians. The word "*like*" in this sentence determines which metaphor academicians' perceptions of AI evoke; in the section continuing with "*because*," the justifications for the metaphors used were revealed. The metaphors were used as the primary data source in the research. Before data collection, the academicians were informed about the research and that the data would be used for only scientific purposes, and they declared their voluntary participation by marking the statement regarding voluntary participation in the data collection tool. Data were collected from academicians working at different universities by sending the form prepared on Google Forms using online platforms due to distance. Data collection permission was obtained from the Batman University Ethics Committee (dated 05.12.2024 and numbered 2024/09-25). The form prepared in the online was kept open to the access of academicians between 9-13th of December. At the end of this period, 104 academicians responded to the online form. However, 11 metaphors were eliminated and excluded from the research for reasons such as incorrect metaphors and the justifications were not meaningful or consistent with the metaphor. Therefore, the raw data in the research consisted of the opinions of 93 academicians. The size of the participant group was determined based on Seidman's (2006) criterion of data saturation. Data saturation is when data is repeated in interviews conducted for research purposes (Seidman, 2006). Thus, the data collection process was terminated when the metaphors repeated.

Data Analysis

The data were analysed using content and descriptive analysis, both widely utilized in qualitative research. Content analysis offers an objective, measurable, and verifiable explanation of the explicit content of messages (Neuendorf, 2017). It also enables a comprehensive examination by categorizing data into themes and uncovering less apparent messages (Krippendorff, 2018). Descriptive analysis complements this by presenting participants' views as they are, with direct quotes reflecting their perspectives without attribution (Yıldırım & Şimşek, 2016).

The analysis followed several stages, as outlined by Saban (2008): coding and sorting, sample metaphor image compilation, category development, and validity and reliability assessment. In the initial coding and sorting stage, the metaphors and justifications provided by participants ($n=104$) were reviewed in detail. Data that lacked meaningful explanations, were irrelevant or failed to establish a connection between metaphor and justification were excluded ($n=11$). The remaining 93 participants were coded anonymously, using the format "A" for academicians followed by a number indicating their participation order (e.g., A-1, A-2, A-3). Personal information was excluded, and participant quotes were presented verbatim to maintain authenticity. In the category development stage, metaphors were grouped into categories and themes based on their justifications. Cohen's Kappa analysis, a statistical measure of inter-rater agreement, was applied to ensure validity and reliability. The agreement score was .78, indicating good consistency in thematic coding and supporting the study's reliability and validity (Cohen, 1960).

FINDINGS

The findings are presented in this section. Academicians used a total of 93 metaphors to describe the concept of AI. Among these, 67 metaphors recurred only once, while 9 metaphors recurred multiple times. Table 1 displays the metaphors, organized by their frequency (f) and corresponding percentage (%).

Table 1. Frequencies (*f*) and percentages (%) of metaphors

No	Metaphor	<i>f</i>	%	No	Metaphor	<i>f</i>	%
1	Child	6	6.45	41	Philosopher	1	1.08
2	Assistant	4	4.30	42	Human intelligence	1	1.08
3	Computer	3	3.23	43	Raincloud	1	1.08
4	Teacher	3	3.23	44	Double-edged sword	1	1.08
5	Human	2	2.15	45	Magnetic ball	1	1.08
6	Friend	2	2.15	46	Safe harbour	1	1.08
7	Guide	2	2.15	47	Heroin	1	1.08
8	Ocean	2	2.15	48	Holy book	1	1.08
9	Universe	2	2.15	49	Reaping without sowing	1	1.08
10	Visor	1	1.08	50	Advisor	1	1.08
11	Medicine	1	1.08	51	Creative robot	1	1.08
12	Knife	1	1.08	52	Angle	1	1.08
13	Air travel	1	1.08	53	Mountain	1	1.08
14	Digital information	1	1.08	54	Autonomous vehicle	1	1.08
15	Headless monster	1	1.08	55	Partially hygienic trash bin	1	1.08
16	Cronus	1	1.08	56	Sky	1	1.08
17	Olympic flame	1	1.08	57	Servant	1	1.08
18	Swimming	1	1.08	58	Devil	1	1.08
19	Memory	1	1.08	59	Genie in the magic lamp	1	1.08
20	Learning process	1	1.08	60	Free cheese	1	1.08
21	Outer space	1	1.08	61	Humanoid robot	1	1.08
22	Tyre	1	1.08	62	Virtual memory	1	1.08
23	Gold bracelet	1	1.08	63	Storage	1	1.08
24	Beyond the human brain	1	1.08	64	Orchestrator	1	1.08
25	Car	1	1.08	65	Factory	1	1.08
26	Vehicle	1	1.08	66	Human brain	1	1.08
27	Ram	1	1.08	67	Mirror	1	1.08
28	Infinity	1	1.08	68	Intelligence	1	1.08
29	Mechanic	1	1.08	69	Tree roots	1	1.08
30	Animal	1	1.08	70	Two sides of the blade	1	1.08
31	Donkey	1	1.08	71	Dream	1	1.08
32	Democracy	1	1.08	72	Mentor	1	1.08
33	Thought factory	1	1.08	73	Library	1	1.08
34	Illusion	1	1.08	74	Grandfather	1	1.08
35	Mind	1	1.08	75	Scale	1	1.08
36	Weapon	1	1.08	76	Motivation booster	1	1.08
40	Password	1	1.08		Total	93	100

As the table shows, the frequency (*f*) of the metaphors ranged from 1 to 6. The most frequently recurred metaphor is "child" (*f*=6; 6.45%), followed by "assistant" (*f*=4; 4.30%), "computer" (*f*=3; 3.23%), and "teacher" (*f*=3; 3.23%), respectively. A total of 67 metaphors were used only once by the academicians. The most frequently recurred metaphors emphasized the positive aspects of AI. However, a closer look at the overall frequency distribution reveals that the negative aspects of AI were also notably emphasized. When the justifications for the metaphors are analysed through content analysis, they are categorized into three main groups: positive, negative, and ambivalent. Metaphors within each category are then explained using themes developed under these three categories.

Positive Category

The metaphors are grouped into four main themes within the positive category based on their common characteristics. These themes are "Humanized Intelligence", "Efficiency and Support", "Learning and Knowledge Transfer", and "Innovation and Creativity". Table 2 presents these themes alongside the metaphors fall under them.

Table 2. Frequencies (f) and percentages (%) of the metaphors under positive categories

Themes	Metaphors	Number of Metaphor	f	%
Humanized Intelligence	Child (4), Human (2), Mirror (1), Human brain (1), Intelligence (1), Autonomous vehicle (1), Philosopher (1), Human intelligence (1), Nervous system (1), Beyond the human brain (1)	10	14	23
Efficiency and Support	Assistant (4), Guide (2), Friend (2), Safe harbour (1), Mind (1), Universe (1), Servant (1), Storage (1), Factory (1), Tree roots (1), Mentor (1), Memory (1), Motivation booster (1), Olympic flame (1), Gold bracelet (1), Virtual memory (1), Air travel (1), Car (1), Password (1), Vehicle (1), Navigation (1), Ram (1), Thought factory (1), Democracy (1)	24	29	48
Learning and Knowledge transfer	Teacher (3), Computer (3), Library (1), Advisor (1), Holy book (1), Grandfather (1), Learning process (1)	7	11	18
Innovation and Creativity	Humanoid robot (2), Creative robot (1), Orchestrator (1), Ocean (1), Mechanic (1), Infinity (1)	6	7	11
Total		47	61	100

As Table 2 shows, a total of 47 distinct positive metaphors were identified. The theme with the most metaphors was *"Efficiency and Support"* (f=29; 48%). This theme was followed by *"Humanized Intelligence"* (f=14; 23%), *"Learning and Knowledge Transfer"* (f=11; 18%), and *"Innovation and Creativity"* (f=7; 11%), respectively.

Humanized Intelligence: Under this theme, academicians used 10 distinct metaphors, representing 23% of the total. This theme emphasizes that AI can think and act like humans and includes features such as AI being perceived as similar to human intelligence and involved in complex thought processes. The metaphors that are most frequently used under this theme were *"child"* (f=4) and *"human"* (f=2).

- Some of the academicians' views representing the theme are as follows:

"AI is like a human because its operating system and artificial neural networks are similar to those of humans." (A-77)

"AI is like a child because it constantly learns and develops." (A-56)

"AI is like a mirror because it can imitate human intelligence." (A-31)

Efficiency and Support: Under this theme, which is under the positive category, academicians used 24 distinct metaphors, representing 48% of the total. The justifications why AI is seen as providing efficiency and support are accelerating work processes, supporting decision-making processes, and saving time. The most frequently recurring metaphors under this theme were *"assistant"* (f=4), *"teacher"* (f=3), *"guide"* (f=2), and *"friend"* (f=2). Based on these findings, it can be concluded that academicians perceive the potential of AI in education and its application areas more in terms of their supportive and productive aspects.

- Some of the academicians' views representing the theme are as follows:

"AI is like a friend because it helps with ideas when needed." (A-32)

"AI is like an assistant because it makes my job easier in daily life and allows me to solve problems that take hours in a short time." (A-48)

"AI is like a guide because it answers questions and guides and makes it easier to access information. However, it leaves all decisions to the user, thus providing the user with a free-thinking and learning space." (A-53)

Knowledge and Learning: Academicians used seven distinct metaphors under this theme, representing 18% of the total. AI's accelerating effect on accessing information and enhancing learning is attributed to its capacity to continuously update knowledge and support teaching processes. The most frequently recurring metaphors under this theme were *"teacher"* (f=3), *"computer"* (f=3), and *"library"* (f=1).

- Some of the academicians' views representing this theme are as follows:

"AI is like a teacher because it guides people to acquire complex knowledge and skills, accelerates the learning process, and contributes to their development." (A-29)

"AI is like a library because it allows me to access the information I am searching for." (A-39)

"AI is like a grandfather because it guides us based on past experiences and leads us to the ideal solution." (A-40)

Innovation and Creativity: Academicians used six distinct metaphors under this theme, accounting for 11% of the total. AI is associated with innovation and creativity because it can generate new ideas and offer creative solutions. The most frequently recurring metaphors under this theme were "creative robot" ($f=1$), "orchestrator" ($f=1$), and "humanized robot" ($f=2$).

- Some of the academicians' views representing this theme are as follows:

"AI is like a creative robot because it can generate ideas and encourage people to develop new solutions." (A-14)

"AI is like an orchestrator because it creates an impressive harmony by blending different knowledge and skills, pushing the boundaries while imitating human intelligence. At the same time, it initiates a journey of thought enriched with emotional and ethical questions." (A-25)

"AI is like a mechanic because you constantly add a new gear, and it gains a new ability to move." (A-91)

Negative Category

Negative metaphors were grouped under four main themes due to their common characteristics. These themes were named "Security Concerns", "Loss of Control", "Misleading", "Ethical Issues and Inequality". Table 3 presents these themes with the metaphors fall under them.

Table 3. Frequencies (f) and percentages (%) of the metaphors under negative categories

Themes	Metaphors	Number of Metaphor	f	%
Security Concerns	Devil (1), Mountain (1), Headless monster (1)	3	3	27
Loss of Control	Heroin (1), Cronus (1), Raincloud (1), Weapon (1), Sky (1)	5	5	45
Misleading	Illusion (1)	1	1	9
Ethical Concerns and Inequality	Free cheese (1), Reaping without sowing (1)	2	2	18
Total		11	11	100

As Table 3 shows, a total of 11 distinct negative metaphors were identified. The theme with the most metaphors was "Loss of Control" ($f=5$; 45%). This theme was followed by "Security Concerns" ($f=3$; 27%), "Ethical Concerns and Inequality" ($f=3$; 18%), and "Misleading" ($f=1$; 9%), respectively.

Security Concerns: Academicians used three distinct metaphors under this theme, accounting for 27% of the total. AI is associated with security concerns due to its potential for malicious use and the risks of losing control over it. The metaphors "devil," "mountain," and "headless monster" stand out in representing this theme.

- Some of the academicians' views representing this theme are as follows:

"AI is like a devil because it is used maliciously." (A-6)

"AI is like a mountain because you can remain under it at any moment." (A-10)

"AI is like a headless monster because there may be some people who are ready to misuse it for their ends." (A-50)

Loss of Control: Academicians had a negative perception of AI due to its potential to take over human control. A total of five distinct metaphors were used under this theme, representing 45% of the total. The metaphors are "heroin," "Cronus" and "weapon." The addictive nature of AI and its uncontrollability are the main justifications for this theme.

- Some of the academicians' views representing this theme are as follows:

"AI is like Cronus because it will eventually swallow its own children in its hands." (A-59)

"AI is like a raincloud because if it rains uncontrollably, it will cause floods and destruction." (A-71)

"AI is like heroin because after it becomes widespread, it will give pleasure and happiness as you use it and will create a great sense of deprivation when you don't use it. It will even meet psychological needs, and people's need for others will decrease even more." (A-18)

Misleading: Concerns about AI's information reliability come to the fore under this theme. Academicians evaluated AI negatively due to the possibility of providing misleading and/or false information. Only one metaphor was used under this theme and represents 9% of the total.

- The academician's view representing the theme is as follows:

"AI is like an illusion because it is thought to work as effectively as the human brain." (A-81)

Ethical Concerns and Inequality: This theme highlights aspects of AI that could violate ethical concerns and create inequality. Academicians used two metaphors under this theme, accounting for 18% of the total. The metaphors representing this theme included *"free cheese"* and *"reaping without sowing"*.

- Some of the academicians' views representing the theme are as follows:

"AI is like free cheese because all information can be obtained with a single click without any effort." (A-4)

"AI is like reaping without sowing because when it is easy to access information, people put in less effort. That is why I see it as a threat to future generations." (A-17)

Ambivalent Category

When the metaphors used by academicians for the concept of AI are examined, it is evident that these metaphors are grouped under three main themes, reflecting both positive and negative characteristics. These themes are named *"Power and Danger," "Innovation and Uncertainty,"* and *"Efficiency and Risk."* In Table 4, these themes are presented along with the metaphors fall under them.

Table 4. Frequencies (f) and percentages (%) of the metaphors under ambivalent category

Themes	Metaphors	Number of Metaphor	f	%
Power and Danger	Scale (1), Medicine (1)	2	2	10
Innovation and Uncertainty	Genie in the magic lamp (1), Partially hygienic trash bin (1), Angle (1), Ocean (1), Universe (1) Dream (1), Magnetic ball (1), Digital information (1), Tyre (1) Outer space (1), Swimming (1) Visor (1)	12	12	57
Efficiency and Risk	Two sides of the blade (1) Animal (1), Double-edged sword (1), Knife (1), Child (2), Donkey (1)	6	7	33
	Total	20	21	100

As Table 4 shows, a total of 20 distinct metaphors were identified. The theme with the most metaphors was *"Innovation and Uncertainty"* (f=12; 57%). This theme was followed by *"Efficiency and Risk"* (f=7; 33%), and *"Power and Danger"* (f=2; 10%), respectively.

Power and Danger: Under this theme there were metaphors that reflected both AI's positive and negative aspects. Academicians used two distinct metaphors, representing 10% of the total. The dual nature of AI, containing both potential power and danger, is conveyed through the metaphors of *"scales"* and *"medicine."* These metaphors suggest that AI can act as a balancing force but may also pose risks when misused.

- Some of the academicians' views representing the theme are as follows:

"AI is like a scale because if you lose the balance of the scale, it can do more harm than good." (A-2)

"AI is like medicine because when used correctly, it can solve problems, but when misused, it can cause harm." (A-65)

Innovation and Uncertainty: This theme focuses on the uncertainties AI entails with its innovative nature. Academicians used 12 distinct metaphors, representing for 57% of this theme. The prominent metaphors included *"genie in the magic lamp," "partially hygienic trash bin," "angel,"* and *"ocean."* These metaphors reflect AI's limitless potential and the unpredictability of its outcomes. The complexity and potential for unexpected results are emphasized.

- Some of the academicians' views representing the theme are as follows:

"AI is like the genie in the magic lamp because it fulfils every wish immediately." (A-5)

"AI is like an angel because it can turn into a devil in the future or remain an angel." (A-13)

"AI is like a visor because it both has and does not have limits. It is integrated with the user's perspective and mental structure." (A-68)

Efficiency and Risk: Under this theme, AI is recognized for its efficiency and inherent risks. Academicians used six distinct metaphors, making up 33% of this theme. Metaphors like *"two sides of the blade," "double-edged sword," "child,"* and *"donkey"* highlight that AI can be beneficial when used properly but risky when misused or uncontrolled.

- Some of the academicians' views representing the theme are as follows:

"AI is like the two sides of a knife because it has positive and negative aspects. While it supports scientific development, it can also lead to problems such as ethical violations." (A-36)

"AI is like the double-edged sword because on the one hand, it is a great invention; on the other hand, it brings with ethical and security concerns." (A-43)

"AI is like a child because if you explain your problem in detail as if you were explaining it to a child, it will easily and surprisingly reflect its imaginary world to you. However, it also has the frightening power to grow and surpass all of us in time." (A-72)

DISCUSSION

This study reveals academicians' mental perceptions of AI through metaphors. The study is significant since it provides insights into how AI is perceived within academia. The findings suggest that perceptions of AI fall into the categories of positive, negative, and ambivalent. Academicians view AI not merely as a technological innovation but also as a concept intertwined with social, academic, and ethical implications.

Educational research provides a basis for understanding the multifaceted perception of AI from an academic perspective by revealing how different educational stakeholders perceive AI. In their study on metaphorical perceptions of AI with three groups of participants, namely students, parents, and teachers, Demirtaş and Türksoy (2023) displayed that participants generally defined AI as inanimate (e.g., robot, machine) and animate entities (e.g., student, brain). Aktaş (2021), in a study conducted with the participation of school administrators and teachers, suggested that participants generally regarded AI as a mechanical and auxiliary entity such as a "robot," "computer," and "assistant." The study concluded that many administrators and teachers perceived AI as a mechanical entity and did not know enough about it. Saçan et al. (2022), in their study examining the metaphorical perceptions of children aged 6-10 towards AI, suggested that they perceived AI as a functioning intelligence, a human-like or robotic entity. It can be concluded that these studies (Aktaş, 2021; Demirtaş & Türksoy, 2023; Saçan et al., 2022) show that metaphors are a powerful tool for understanding the role and perception of AI in education.

The current study investigated academicians' perceptions of AI through metaphors and revealed under which categories the perceptions could be grouped. The findings show that academicians address AI in various aspects and that these evaluations are concentrated under specific main themes. Considering the themes that include positive, negative, and ambivalent characteristics, the most metaphors were used under the theme of "Efficiency and Support." This theme highlights features such as AI's acceleration of work processes, support for decision-making processes, and time-saving, which previous research frequently emphasized (Brynjolfsson & McAfee, 2014; West, 2018). This finding is also consistent with the previous research (Aktaş, 2021; Demirtaş & Türksoy, 2023). These studies also suggested that education stakeholders generally perceived AI as a helpful and efficient tool. On the other hand, the fact that AI was regarded as a mechanical entity in these studies showed that the participants evaluated AI as a technical support tool. The theme of "Humanized Intelligence" refers to the potential of AI to think and operate like human intelligence. In this theme, it is noteworthy that AI is likened to learning and development processes. The theme of "Knowledge and Learning" marks the capacity of AI to provide ease of access to information and to support learning processes. This theme shows that AI is defined by "teacher", "library", and "computer" metaphors. The theme of "Innovation and Creativity" emphasizes AI's innovative and creative aspects. In this theme, the capacity of AI to generate creative solutions and the ability to adapt by bringing different elements together come to the fore.

When negative metaphors are analyzed, the "Security Concerns" theme highlights the potential risks associated with AI in case of its uncontrolled use. Academicians emphasized that improper or uncontrolled use of AI could have harmful consequences, a concern echoed in an international report (OECD, 2024). The theme of "Loss of Control" stems from concerns that AI could overtake human autonomy and foster dependency or addiction. The theme of "Misleading" underscores apprehensions about the reliability of AI-generated information and the adverse outcomes of failing to differentiate between accurate and false information. Similarly, the theme of "Ethical Issues and Inequality" raises questions about AI's capacity to challenge ethical norms and exacerbate disparities. These ethical and inequality concerns align with existing literature, including studies by Abrams (2024) and Jobin et al. (2019). Additionally, these findings resonate with Saçan et al.'s (2022) conclusion that children often perceive AI as an entity requiring control. Furthermore, Nartgün and Kennedy's (2024a) exploration of AI-related educational challenges corroborates academicians' ethical and security concerns. However, the depth of scrutiny academicians apply to AI's ethical and societal implications suggests that their perspectives, shaped by adult and academic experiences, are more critical and reflective than those of children and other stakeholders.

The ambivalent themes suggest that AI has a bilateral nature. The theme of "Power and Danger" implies that AI can potentially be both beneficial and dangerous. The theme of "Innovation and Uncertainty" emphasizes the innovative nature of AI and the uncertainties it brings about. This theme indicates that academicians have bilateral perceptions of the limitless potential and unpredictable results of AI. The theme of "Efficiency and Risk" emphasizes that AI can provide efficiency but shelter risks. This theme states that AI can be beneficial when used correctly, but it shelter risks when used uncontrollably or unconsciously. These findings are consistent with the "Guide Dog Metaphor" of Boulard et al. (2023), which considers AI as both a guiding and controllable entity. While this study emphasizes that AI is steerable and user-friendly, it is understood that academicians are also concerned about the unpredictable aspects of AI.

The current research reveals that academicians evaluate AI with both potential benefits and risks. In particular, the views of academicians on the need to be careful and attentive about the ethical use of AI and its effects on society provide an important perspective that should be taken into account in the process of developing AI policies and practices. The fact that AI is regarded as a tool that facilitates human life accelerates access to information and supports creative processes shows that it can provide social benefits when used properly. However, it is thought that preventing uncontrolled use and ethical violations will ensure a safer and more effective use of AI technologies.

CONCLUSION AND RECOMMENDATIONS

The findings showed that metaphors fell within three categories which were positive, negative and ambivalence. In the positive category “*Humanized Intelligence, Efficiency and Support, Learning and Knowledge Transfer, Innovation and Creativity*”; in the negative category “*Security Concerns, Loss of Control, Misleading, Ethical Concerns and Inequality*” and ambivalence category “*Power and Danger*”, “*Innovation and Uncertainty*”, “*Efficiency and Risk*” were the emerging themes. The most metaphors were used under the theme of “*Efficiency and Support*”. This study reveals that academicians generally perceive AI positively as a tool accelerating work processes and increasing efficiency. However, AI has negative aspects, such as security concerns, loss of control, ethical issues, and inequality. The study reveals that academicians tend to make a balanced assessment of AI, carefully weighing its potential benefits against its associated risks. This nuanced perspective offers valuable insights that can inform the development of future policies and best practices regarding AI integration.

Based on the findings of the current study, the researchers bring the following recommendations. Training programs can be held for academicians, researchers, practitioners, and policymakers to raise awareness about AI's academic and social impacts. These programs should provide information on the proper use of AI, ethical standards, and risk management which can maximize the potential of AI in academia and education. Further research can provide a broader framework by including the views of academicians from different countries and incorporating quantitative methods. In addition, addressing AI's academic and societal impacts with a holistic approach can be a substantial direction for future research and practices.

LIMITATIONS

The current study is without limitations. First, it was limited to the views of academicians working in Türkiye. Second, it adopted a qualitative approach, and the data were collected through metaphors. Although this provides an in-depth view, it limits the generalizability of the findings. Lastly, focusing only on the views of academicians limited the opportunity to provide a perspective on other stakeholders' perceptions of AI.

Declaration of Conflicting Interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The authors received no financial support for the research, author-ship, and/or publication of this article.

Statements of publication ethics

We hereby declare that the study has not unethical issues and that research and publication ethics have been observed carefully.

Researchers' contribution rate

The study was conducted and reported with equal collaboration of the researchers.

Ethics Committee Approval Information

Batman University Ethics Committee (dated 05.12.2024 and numbered 2024/09-25).

REFERENCES

- Abrams, Z. (2024). Addressing equity and ethics in artificial intelligence. *Monitor on Psychology*, 55(3), 1-24.
- Akbulak, Y. (2022). Avrupa Birliği genel veri koruma yönetmeliğinin temel esasları. *Legal Blog*. Erişim adresi: <https://legal.com.tr/blog/genel/avrupa-birligi-genel-veri-koruma-yonetmeliginin-temel-esaslari/>
- Aktaş, A. (2021). Yönetici ve öğretmen görüşlerine göre yapay zekâ: Bir metafor çalışması. 1. *Ulusal Eğitimde Yapay Zekâ Uygulamaları Kongresi Tam Metin Bildiri Kitapçığı* içinde (ss. 4-36) Şanlıurfa, Türkiye.
- Altıntop, M. (2023). Yapay zekâ/akıllı öğrenme teknolojileriyle akademik metin yazma: ChatGPT örneği. *Süleyman Demirel Üniversitesi Sosyal Bilimler Enstitüsü Dergisi*, 46(2), 186-211. <https://doi.org/10.3386/w24285>
- Ateş, F. (2024). Yapay zekânın tarihi. *Teknodestek*. Erişim adresi: <https://teknodestek.com.tr/yapay-zekanin-tarihi/>
- Ateşgöz, K. (2023). Otomatikleşen medya ortamı bağlamında yapay zekâ haberciliği/gazeteciliği okuryazarlığı. İçinde *Medya ve Habercilik Alanında Yapay Zekânın Yükselişi* (Edt.: K. Ateşgöz) (ss. 95-116). Ankara: Eğitim Yayınevi.
- Aubusson, P. (2002). Using metaphor to make sense and build theory in qualitative analysis. *The Qualitative Report*, 7(4), 1-14. <https://doi.org/10.46743/2160-3715/2002.1962>
- Barocas, S., Hardt, M., & Narayanan, A. (2018). Fairness and machine learning. Erişim adresi: <https://fairmlbook.org/>
- Boulard, C., Viswanathan, S., & Jacquin, T. (2023). The guide dog metaphor: Explainable AI via coachable AI. *Proceedings of the ACM CHI Conference on Human Factors in Computing Systems*. <https://doi.org/10.13140/RG.2.2.21966.74560>
- Brynjolfsson, E., & McAfee, A. (2014). *The second machine age: Work, progress, and prosperity in a time of brilliant technologies*. WW Norton & Company.
- Castells, M. (2011). *The rise of the network society*. John Wiley & Sons.

- Chiu, T. K. F., Xia, Q., Zhou, X., Chai, C. S., & Cheng, M. (2023). Systematic literature review on opportunities, challenges, and future research recommendations of artificial intelligence in education. *Computers and Education: Artificial Intelligence*, 4: 100118, 1-15. <https://doi.org/10.1016/j.caeai.2022.100118>
- Chui, M., Manyika, J., & Miremadi, M. (2016). *Where machines could replace humans-and where they can't (yet)*. www.mckinsey.com
- Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, 20(1), 37-46. <https://doi.org/10.1177/001316446002000104>
- Copeland, B. J. (2024). History of artificial intelligence (AI). In *Encyclopedia Britannica*. Retrieved November 5, 2024, from <https://www.britannica.com/science/history-of-artificial-intelligence>
- Crompton, H., & Burke, D. (2023). Artificial intelligence in higher education: The state of the field. *International Journal of Educational Technology in Higher Education*, 20, 1-22. <https://doi.org/10.1186/s41239-023-00392-8>
- Crompton, H., & Song, D. (2021). The potential of artificial intelligence in higher education. *Revista Virtual Universidad Católica Del Norte*, 62, 1-4. <https://doi.org/10.35575/rvucn.n62a1>
- Demirtaş, E., & Türksoy, Ü. E. (2023). İlkokul eğitim paydaşlarının yapay zekâya yönelik metaforik algıları. *Tam Metinler Kitabı Ufuk University 2 Nd International Congress on Social Sciences* (ss. 194-207).
- Dolunay, A., & Temel, A. C. (2024). The relationship between personal and professional goals and emotional state in academia: A study on unethical use of artificial intelligence. *Frontiers in Psychology*, 15:1363174, 1-10. <https://doi.org/10.3389/fpsyg.2024.1363174>
- Ece, N. J. (2024). Yapay zeka: Denizcilik sektöründe kullanımı ve SWOT analizi. *Mersin Üniversitesi Denizcilik ve Lojistik Araştırmaları Dergisi*, 6(1), 30-53. <https://doi.org/10.54410/denlojad.1491372>
- Efe, A. (2022). Yargısal ve hukuki süreçlerde yapay zekâ kullanan araçlar üzerine bir araştırma. *Bilgi Yönetimi*, 5(1), 92-117. <https://doi.org/10.33721/by.971873>
- Eryılmaz, H. E. (2023). Yapay zekâ çağında kişisel veri mahremiyeti. *UMAY Sanat ve Sosyal Bilimler Dergisi*, 1(2), 6-25. <https://doi.org/10.55709/TSBSBildirilerDergisi.539>
- Etike, Ş. (2023). Yapay zekâyı insan ve kamusal yarar için çalıştırmak: ABD ve AB politika belgelerinin önerileri. *Memleket Siyaset Yönetim*, 18(39), 220-249. <https://doi.org/10.56524/msydergi.1301819>
- Farahani, M. S., & Ghasemi, G. (2024). Artificial intelligence and inequality: Challenges and opportunities. *Int. J. Innov. Educ*, 9, 78-99. <https://doi.org/10.32388/7HWU22>
- Görgülü, D., Coşkun, F., Demir, M., & Sipahioğlu, M. (2025). A psychometric analysis of the artificial intelligence skills scale developed through chat gpt. *Education and Information Technologies*. <https://doi.org/10.1007/s10639-024-13294-7>
- Güven, H., & Ayvaz Güven, E. T. (2023). Yapay zekâ uygulamalarının e-ticarette kullanımı. *International Journal of Management and Administration*, 7(13), 69-94. <https://doi.org/10.29064/ijma.1194949>
- Haenlein, M., & Kaplan, A. (2019). A brief history of artificial intelligence: On the past, present, and future of artificial intelligence. *California Management Review*, 61(4), 5-14. <https://doi.org/10.1177/0008125619864925>
- Holmes, W., Bialik, M., & Fadel, C. (2019). *Artificial intelligence in education: Promises and implications for teaching and learning*. Boston: Center for Curriculum Redesign.
- Huang, C., Zhang, Z., Mao, B., & Yao, X. (2023). An overview of artificial intelligence ethics. *IEEE Transactions on Artificial Intelligence*, 4(4), 799-812. <https://doi.org/10.1109/TAI.2022.3194503>
- Jobin, A., Ienca, M., & Vayena, E. (2019). The global landscape of AI ethics guidelines. *Nat Mach Intell*, 1, 389-399. <https://doi.org/10.1038/s42256-019-0088-2>
- Joyce, K., Smith-Doerr, L., Alegria, S., Bell, S., Cruz, T., Hoffman, S. G., Noble, S. U., & Shestakofsky, B. (2021). Toward a sociology of artificial intelligence: A call for research on inequalities and structural change. *Socius*, 7(4), 1-11. <https://doi.org/10.1177/2378023121999581>
- Kaplan, A., & Haenlein, M. (2020). Rulers of the world, unite! The challenges and opportunities of artificial intelligence. *Business Horizons*, 63(1), 37-50. <https://doi.org/10.1016/j.bushor.2019.09.003>
- Keleş, H. (2022). Tıpta yapay zeka uygulamaları. *Kırıkkale Üniversitesi Tıp Fakültesi Dergisi*, 24(3), 604-613. <https://doi.org/10.24938/kutfd.1214512>
- Kırık, A. M., & Özkoçak, V. (2023). Medya ve iletişim bağlamında yapay zekâ tarihi ve teknolojisi: ChatGPT ve Deepfake ile gelen dijital dönüşüm. *Karadeniz Uluslararası Bilimsel Dergi*, 58, 73-99. <https://doi.org/10.17498/kdeniz.1308471>
- Krippendorff, K. (2018). *Content analysis: An introduction to its methodology*. Sage Publications. <https://doi.org/10.4135/9781071878781>
- Lakoff, G., & Johnson, M. (1981). *Metaphors we live by*. University of Chicago Press. <https://doi.org/10.7208/chicago/9780226470993.001.0001>
- Li, G., Deng, X., Gao, Z., & Chen, F. (2019). Analysis on ethical problems of artificial intelligence technology. In *Proceedings of the 2019 International Conference on Modern Educational Technology* (pp. 101-105). <https://doi.org/10.1145/3341042.3341057>
- Livberber, T., & Ayvaz, S. (2023). The impact of Artificial Intelligence in academia: Views of Turkish academics on ChatGPT. *Heliyon*, 9(9):e19688, 1-11. <https://doi.org/10.1016/j.heliyon.2023.e19688>
- Midgley, W., Trimmer, K., & Davies, A. (Eds.). (2014). *Metaphors for, in and of education research*. Cambridge Scholars Publishing.
- Mijwil, M. M., Hiran, K. K., Doshi, R., Dadhich, M., Al-Mistarehi, A. H., & Bala, I. (2023). ChatGPT and the future of academic integrity in the artificial intelligence era: A new frontier. *Al-Salam Journal of Engineering and Technology*, 2, 116-127. <https://doi.org/10.55145/ajest.2023.02.02.015>
- Mijwil, M. M., Sadıkoğlu, E., Cengiz, E., & Candan, H. (2022). Siber güvenlikte yapay zekanın rolü ve önemi: Bir derleme. *Veri Bilimi*, 5(2), 97-105.
- Mitchell, T. M. (1997). *Machine learning*. McGraw-Hill.
- Nartgün, Z., & Kennedy, E. (2024a). Artificial Intelligence (AI) and education: Contributions opportunities and challenges. *Journal of Economic and Politic Studies in Education*, 5(2), 1-23.
- Nartgün, Z., & Kennedy, E. (2024b). Cheating in higher education in the age of artificial intelligence. *International Journal on Lifelong Education and Leadership*, 10(2), 47-53. <https://doi.org/10.25233/ijlel.1553831>
- Neuendorf, K. A. (2017). *The content analysis guidebook*. Thousand Oaks, CA: Sage. <https://doi.org/10.4135/9781071802878>
- Niskanen, T., Sipola, T., & Väänänen, O. (2023). Latest trends in artificial intelligence technology: A scoping review. *Jamk University of Applied Sciences*. <https://arxiv.org/html/2305.04532>

- Noy, S., & Zhang, W. (2023). Experimental evidence on the productivity effects of generative artificial intelligence. *Science*, 381(6654), 187-192. <https://doi.org/10.1126/science.adh2586>
- OECD (2024). *Assessing potential future artificial intelligence risks, benefits and policy imperatives*. OECD Artificial Intelligence Papers, No. 27. <https://oecd.ai/site/ai-futures>
- Özdemir, A. (2023). Finans sektörünü yapay zekâ ile birlikte okumak: Yenilikler, fırsatlar ve engeller. İçinde *Dijitalleşmenin Finans Sektörüne Getirdiği Yenilikler* (Edt.: R. Yücel, Y. Ayyıldız, & H. Er), (pp. 57-70). Özgür Yayınları. <https://doi.org/10.58830/ozgur.pub298.c1468>
- Par, S. (2024). Yapay zekânın sağlık alanında kullanımı ve hukuki statüsü. *International Anatolia Academic Online Journal Social Sciences Journal*, 10(2), 179-196.
- Pedro, F., Subosa, M., Rivas, A., & Valverde, P. (2019). *Artificial intelligence in education: Challenges and opportunities for sustainable development*. <https://unesdoc.unesco.org/ark:/48223/pf0000366994>
- Polat, M. (2024). Dijitalleşme ve eğitimde yapay zekâ: Eğitim yönetimine yansımaları. *Education Science and Sports*, 6(2), 1-12. <https://doi.org/10.70053/esas.1566395>
- Russell, S., & Norvig, P. (2020). *Artificial intelligence: A modern approach*. (4th ed.). Pearson.
- Saban, A. (2008). Okula ilişkin metaforlar. *Kuram ve Uygulamada Eğitim Yönetimi*, 55, 459-496.
- Saçan, S., Tozduman Yaralı, K., & Kavruk, S. Z. (2022). Çocukların "yapay zekâ" kavramına ilişkin metaforik algılarının incelenmesi. *Mehmet Akif Ersoy Üniversitesi Eğitim Fakültesi Dergisi*, 64, 274-296. <https://doi.org/10.21764/maeuefd.1074024>
- Sariyasa, S., & Monika, K. A. L. (2023). Artificial intelligence and academic ethics in the era of Merdeka Belajar: How are students' responses? *Jurnal Kependidikan: Jurnal Hasil Penelitian dan Kajian Kepustakaan di Bidang Pendidikan, Pengajaran dan Pembelajaran*, 9, 986-995. <https://doi.org/10.33394/jk.v9i3.8720>
- Schmitt, R. (2005). Systematic metaphor analysis as a method of qualitative research. *The Qualitative Report*, 10(2), 358-394. <https://doi.org/10.46743/2160-3715/2005.1854>
- Schultz, D. P., & Ellen-Schultz, S. (2007). *Modern psikoloji tarihi* (Çev.: Y. Aslay). İstanbul: Kaknüs Yayınları.
- Schwab, K. (2017). *The fourth industrial revolution*. Crown Business.
- Seidman, I. (2006). *Interviewing as qualitative research: A guide for researchers in education and the social sciences* (3rd ed.). New York: Teachers College Press.
- Sheikhi, M. (2022). Yapay zeka kullanımının iş piyasasına etkisi. *Journal of Economics and Political Sciences*, 2(1), 102-111.
- Sipahioğlu, M. (2022). Toplum 5.0 ve eğitim örgütlerinde dijital dönüşüm. İçinde *Eğitim Kurumlarında Örgütsel Davranış* (Edt.: S. Güven), ss. 15-40. Ankara: Nobel Akademik Yayıncılık.
- Sipahioğlu, M. (2023). Sınıfta teknoloji yönetimi. İçinde *Sınıf Yönetiminde Çağdaş Yaklaşımlar* (Edt.: D. Görgülü ve R. Küçükali). Ankara: Nobel Akademik Yayıncılık.
- Şen, F., & Gür, B. (2024). Dış ticaret süreçlerinde yeni nesil teknolojilerin kullanılmasına dair kavramsal bir çerçeve. *Eurasian Academy of Sciences Social Sciences Journal*, 52, 66-79. <http://doi.org/10.17740/eas.soc.2024.V52.04>
- Thunström, A. O. (2022). We asked GPT-3 to write an academic paper about itself- then we tried to get it published: an artificially intelligent first author presents many ethical questions-and could upend the publishing process. *Scientific American*, 327(3), 70. <https://doi.org/10.1038/scientificamerican0922-70>
- Turing, A. M. (1950). Computing machinery and intelligence. *Mind*, LIX(236), 433-460. <https://doi.org/10.1093/mind/LIX.236.433>
- WEF (2020). *The Future of jobs report 2020*. Geneva: World Economic Forum. Erişim adresi: <https://www.weforum.org/publications/the-future-of-jobs-report-2020/>
- West, D. M. (2018). *The future of work: Robots, AI, and automation*. Brookings Institution Press.
- Yenikaya, M. A., & Oktaysoy, O. (2023). Yapay zekâ uygulamalarının sağlık sektöründe kullanımı: Derin öğrenme yöntemiyle ön tanı. *Sakarya Üniversitesi İşletme Enstitüsü Dergisi*, 5(2), 127-131. <https://doi.org/10.47542/sauied.1394746>
- Yıldırım, A., & Şimşek, H. (2016). *Sosyal bilimlerde nitel araştırma yöntemleri*. Ankara: Seçkin Yayıncılık.
- Yıldız, A. (2022). Finans alanında yapay zekâ teknolojisinin kullanımı: Sistematik literatür incelemesi. *Pamukkale Üniversitesi Sosyal Bilimler Enstitüsü Dergisi*, 52, 47-66.
- Yiğit, S., Berşe, S., & Dirgar, E. (2023). Yapay zekâ destekli dil işleme teknolojisi olan ChatGPT'nin sağlık hizmetlerinde kullanımı. *Eurasian Journal of Health Technology Assessment*, 7(1), 57-65. <https://doi.org/10.52148/ehta.1302000>
- Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education – where are the educators? *International Journal of Educational Technology in Higher Education*, 16, 39. <https://doi.org/10.1186/s41239-019-0171-0>
- Zlateva, P., Steshina, L., Petukhov, I., & Velev, D. (2024). A conceptual framework for solving ethical issues in generative AI. In *Electronics, Communications and Networks* (pp. 110-119). IOS Press. <https://doi.org/10.3233/FAIA231182>



Navigating the Accreditation Labyrinth with Meta Synthesis Tips

Meta Sentez İpuçları İle Akreditasyon Labirentinde Yol Bulmak

Alev ORHAN¹

Keywords

1. Accreditation
2. Meta-synthesis
3. Higher education
4. Quality of higher education

Anahtar Kelimeler

1. Akreditasyon
2. Meta sentez
3. Yükseköğretim
4. Yükseköğretimin niteliği

Received/Başvuru Tarihi

24.03.2025

Accepted / Kabul Tarihi

10.10.2025

Abstract

Purpose: This approach will facilitate the identification of prevailing issues and recent developments in the accreditation process. The objective of this study is to systematically analyze and present all aspects of qualitative theses that address the accreditation process. Furthermore, the objective was to ascertain the lacunae within the field by employing meta-synthesis analysis, a methodological approach that facilitates the examination of all dimensions of qualitative research on the accreditation process.

Design/Methodology/Approach: In this study, meta-synthesis, a qualitative research method, was employed. The present study encompasses the qualitative findings of a set of graduate theses (n = 17) obtained from the Thesis Center of the Council of Higher Education. A content analysis was performed on the data obtained through the utilization of the MAXQDA 2020 (20.4.0) program.

Findings: The results of the research demonstrated that the majority of accreditation studies are conducted within the domain of educational sciences, with a preponderance of literature focusing on the comparison of accreditation processes across various nations. Furthermore, the research findings indicate that the impact of training curricula on accreditation is a salient issue, underscoring the significance of accreditation by independent institutions.

Highlights: Researchers, too, have several recommendations, including the inclusion of different stakeholders.

Öz

Çalışmanın amacı: Yükseköğretim, öğretim, bilimsel araştırma ve topluma hizmet işlevlerinin gerçekleştirildiği dinamik bir süreçtir. Bu noktada, mevcut literatürün kapsamlı ve bütüncül bir analizinin yapılması zorunludur. Bu analiz, ülkelerin akreditasyon kriterleri, akreditasyon kurumlarının özellikleri, öğretim süreci ve kurumları ile ilgili değerlendirme ve karşılaştırmaları kapsamalıdır. Bu yaklaşım, akreditasyon sürecindeki mevcut sorunların ve son gelişmelerin belirlenmesini kolaylaştıracaktır. Bu çalışmanın amacı, akreditasyon sürecini ele alan nitel tezlerin tüm yönlerini sistematik olarak analiz etmek ve sunmaktır. Ayrıca, akreditasyon sürecine ilişkin nitel araştırmaların tüm boyutlarının incelenmesini kolaylaştıran metodolojik bir yaklaşım olan meta-sentez analizi kullanılarak alandaki boşlukların tespit edilmesi amaçlanmıştır.

Materyal ve Yöntem: Bu çalışmada nitel araştırma yöntemlerinden olan meta-sentez kullanılmıştır. Veri seti tez tarama veri tabanından “akreditasyon” anahtar kavramı kullanılarak elde edilmiştir. Araştırmanın kapsamını ise Yükseköğretim Kurulu Başkanlığı Tez Merkezi’nden elde edilen lisansüstü tezlerin (n=17) nitel bulguları oluşturmaktadır. Elde edilen veriler üzerinden MAXQDA 2020 (20.4.0) programı aracılığıyla içerik analizi gerçekleştirilerek ilgili tema ve kodlar oluşturulmuştur.

Bulgular: Araştırma sonuçları akreditasyon çalışmalarının en fazla eğitim bilimleri alanında yapıldığını ve ülkeler arasındaki akreditasyon sürecinin karşılaştırılmasının daha ağırlıklı olduğunu göstermiştir. Ayrıca araştırma sonuçlarına göre eğitim müfredatlarının akreditasyon üzerindeki etkilerinin ön plana çıktığı ve akreditasyonun bağımsız kurumlar tarafından yapılmasının önemine dikkat çekildiği görülmektedir. İncelenen tezlerden ulaşılan bir başka çarpıcı sonuç ise teknik altyapının akreditasyonla ilişkilendirilmesidir.

Önemli Vurgular: İncelenen tezlerin önerileri dikkate alındığında ise uygulayıcılara yönelik en fazla akreditasyon için öğrenci biriminin kurulması, araştırmacılara yönelik ise programların disiplin alanlarına yönelik ülke karşılaştırmalarının yapılması ve farklı paydaşlara yer verilmesi önerileri daha fazla vurgulanmıştır.

¹ Sivas Cumhuriyet University, Yıldızeli Vocational School, Sivas, Türkiye; <https://orcid.org/0000-0002-8999-9329>

INTRODUCTION

It can be posited that quality management has existed in industry and the public sector for approximately a century (Lagrosen, 2017). The history of quality management in higher education is relatively recent in comparison to that of industry and the public sector. However, it has gained significant importance in recent years. Higher education is a dynamic and complex system, comprising a variety of activities and processes. The higher education system can be described in terms of an input, transformation, and process model. Inputs in the context of higher education can be broadly defined as comprising human, physical, and financial resources. The outputs of this research are graduates, research findings, and service to society (Mati, 2018). There are ongoing efforts to achieve accreditation in order to ensure the continued provision of higher education in a manner that is both qualified and functional, with a focus on the core functions of education, scientific research, and service to society. In consequence of technological advancement, the proliferation of distance learning and the globalisation of higher education, educational institutions and programmes require a dependable system to guarantee quality. It can be argued that an essential aspect of quality management in higher education is accreditation.

Accreditation can be conceptualized as a mechanism through which educational institutions, including universities and academic programs, can substantiate the quality of their offerings to both students and the general public (Eaton, 2015). From an alternative standpoint, accreditation may be regarded as a pledge to guarantee quality and to uphold the equitable framework of higher education (López Segrera, Sanyal & Tres, 2007). The accreditation of higher education institutions is a significant factor in determining their prestige and appearance, as it is widely regarded as an important indicator of future success. In light of these considerations, there is a growing tendency to make accreditation a mandatory program, rather than a voluntary one (Nicklin, Engel & Stewart, 2021). Accreditation is designed as a tool to promote and facilitate the achievement of international quality (Sánchez-Chaparro, Remaud, Gómez-Frías, Duykaerts & Jolly, 2022) and is aligned with the institution's mission, vision, impact, and innovation. The institution's intellectual contribution, impact, and mission; students' progression and career development; the competence and assignments of teaching staff; curriculum arrangements and content; student-instructor interactions; administrator qualifications; and instructor qualifications were identified as accreditation criteria (Lagrosen, 2017). While the specific criteria may vary depending on the discipline for which accreditation is sought, the general framework of accreditation standards remains consistent across different fields (Mati, 2018). The objective of these criteria is to enhance the sustainability and quality of higher education programs (Ebekozién & Aigbavboa, 2023). In order to maintain quality and facilitate the emergence of academic leadership, it is essential to guarantee academic autonomy (Eaton, 2015). This is regarded as a crucial measure towards enhancing transparency and accountability within the European quality system. In order to guarantee objectivity and accountability, internal and external quality assurance systems have been implemented. These systems guarantee that students and graduates can participate in educational systems across countries without impediment (Sánchez-Chaparro et al., 2022). It is imperative that accreditation processes and criteria are aligned to ensure uniformity in educational quality. Primacy should be accorded to the institutional vision in the accreditation evaluation process (Eaton, 2015). A subsequent crucial phase is the establishment of learning objectives. These objectives must be congruent with the institution's vision, learning outcomes, and methodologies (Kundu & Majumdar, 2020). In this process, the significance of the curriculum is underscored, and it is asserted that the absence of an optimal curriculum design has a detrimental impact on the quality of education (Fernandes & Singh, 2022). Consequently, when developing curricula, it is essential to ensure that the team comprises representatives from various stakeholder groups, including alumni, student organizations, university administrators, faculty members, and that the curriculum is aligned with the established learning goals (Kundu & Majumdar, 2020). An essential element in the accreditation process is the quality of the teaching staff and the physical and technical infrastructure necessary for qualified teaching (Asiyai, 2015). At this juncture, instructors are tasked with ensuring efficacious student-student and student-faculty interaction. They are also encouraged to leverage this interaction to facilitate collaborative learning opportunities (Kundu & Majumdar, 2020).

The primary distinction that sets this study apart from analogous studies is the systematic meta-synthesis of graduate theses on the accreditation process. Despite the existence of numerous studies addressing accreditation in the extant literature, the preponderance of these studies has been confined to specific contexts, thus failing to offer a comprehensive framework. However, the present study exclusively encompasses qualitative and mixed-methods theses, with a comprehensive analysis conducted on multidimensional variables. These include publication type, publication year, institution, discipline, method, design, sample, data collection tool, data analysis technique, credibility and consistency, research problems, and research results. Furthermore, the study's examination of the dimensions of credibility, consistency, and objectivity makes a unique contribution in terms of providing a critical assessment of methodological quality. In these respects, the study makes original contributions to the field at both theoretical and practical levels. Consequently, the study offers theoretical contributions as well as concrete implications for practitioners and suggestions for future research. Moreover, the evaluation of theses from disparate institutions and contexts facilitates a comparative analysis of prevalent trends, problem areas, and recommendations for enhancing accreditation processes. A comprehensive review of qualitative theses concerning the accreditation process was conducted, resulting in the identification of its strengths and weaknesses. Furthermore, this analysis revealed existing gaps in the field and offered guidance to practitioners. Consequently, it is imperative in terms of providing scientific contribution to both future academic studies and stakeholders involved in the accreditation process.

RESEARCH PROBLEM: A critical examination of qualitative and mixed-method graduate theses pertaining to the accreditation process is imperative to ascertain the extent to which these documents address all aspects of the process. Furthermore, it is

essential to elucidate the findings revealed by these theses concerning the problems, suggestions, and practices encountered during the process.

SUB-PROBLEMS

1. What is the distribution of qualitative theses in terms of publication type, year of publication, universities and disciplinary fields?
2. What are the research methods, designs and sample types used in the theses analyzed?
3. In which ways were the data collection tools and data analysis techniques used in the theses applied?
4. What are the findings on how credibility and consistency are ensured in theses?
5. Which topics do the research problems addressed in the theses focus on?
6. What findings were found in the research results about the accreditation process and what kind of recommendations were made?

METHOD/MATERIALS

In this study, a meta-synthesis, one of the qualitative research methods, was employed. A meta-synthesis entails the qualitative aggregation and interpretation of non-quantitative findings through the utilization of diverse methodologies. The principal objective of numerous meta-synthesis studies is to devise and refine disparate qualitative findings in a meaningful manner (Finfgeld-Connett, 2010). It is anticipated that this design will facilitate the synthesis of qualitative research findings on the concept of accreditation, the identification of problematic areas, and the formulation of recommendations based on the examination of relevant studies. The synthesis of these findings will inform future practices and research in this field.

Data Collection Process and Inclusion Criteria

In the study, criterion sampling was the preferred method of data collection. Criterion sampling is a preferred sampling method in both qualitative and quantitative research. The principal objective of criterion sampling is to conduct in-depth investigations according to pre-established criteria that encompass significant attributes of the event or phenomenon under study (Patton, 2002). The research included studies related to accreditation. Initially, the subject tab was populated with the term "education and training," and the keyword tab was populated with the term "accreditation." This was done to ensure that studies pertinent to the research topic were identified. The inclusion of postgraduate theses in the study was based on the criteria of qualitative or mixed research methods and the inclusion of the concept of "accreditation" in the keywords. On September 2, 2024, a total of 33 postgraduate theses were deemed eligible for inclusion in the study, with 16 being excluded due to their design employing quantitative methods. Following the application of the established criteria, 17 theses were ultimately selected for analysis. A publication review form was devised by the researcher on September 27, 2024, with the objective of examining the data obtained. In order to ascertain the general characteristics of the studies, the titles of the type of publication, the year of publication, the university where the thesis was conducted, and the disciplinary field were included. Furthermore, in order to elucidate the interconnections within meta-synthesis reviews, a comprehensive array of methodological elements was considered, including the methodology, sampling strategy, study group, data collection instrument, analytical techniques, credibility and consistency, research issues, and outcomes. The draft form was submitted to a researcher with expertise in the field of educational sciences for the purpose of conducting a meta-synthesis analysis. In light of the expert feedback, the feedback form was divided into four sections: credibility and consistency, objectivity, and suggestions.

The Validity and Reliability of Research

In selecting the 17 postgraduate theses comprising the data set for the study, those that were accessible were included in the research using the YÖK thesis screening infrastructure and referenced accordingly. At this juncture, efforts were undertaken to guarantee that the reader could gain access to the data set. To guarantee the veracity and reliability of the study, all processes involved in data coding and analysis were elucidated in exhaustive detail and subjected to rigorous scrutiny by the researcher using the MAXQDA program. To guarantee the dependability of the findings, the publication review form and coding list were submitted to an expert in the field of educational sciences for evaluation. Based on the expert's feedback, the necessary corrections were implemented.

Data Analysis

A content analysis method was employed to analyze the data obtained from the publication review form, which was prepared for the purpose of addressing the research question. The theses that were identified within the scope of the research were assigned a code (T1, T2, etc.) and entered into a file. The research data were subjected to analysis using the MAXQDA 2020 (20.4.0) program, with the generation of themes and codes.

FINDINGS

Following the completion of the evaluation process, 17 postgraduate theses included in the research project were selected for inclusion in the descriptive analysis. A descriptive analysis of the analyzed studies is presented in Figure 1.

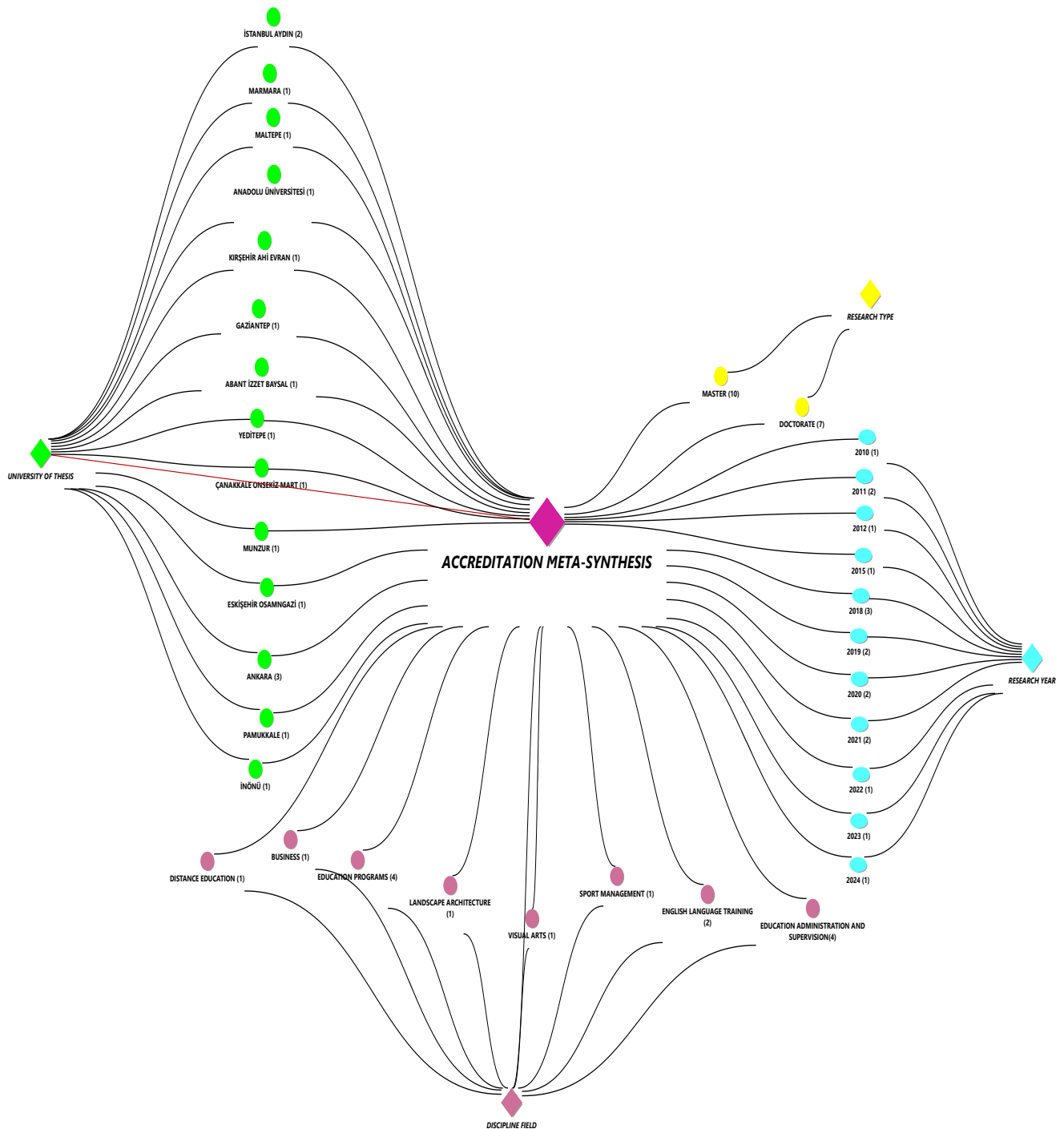


Figure 1. Code co-creation Model on Accreditation Meta-Synthesis

Upon examination of Figure 1, it becomes evident that the inaugural thesis study pertaining to the concept of accreditation was conducted in 2010 and subsequently gained momentum in 2018. From 2018 to the present day, it can be stated that the research output has demonstrated a decline in momentum. A review of the literature reveals that the majority of research studies are master's theses (10). A review of the literature revealed that Ankara University and Istanbul Aydın University have conducted a total of five theses on the concept of accreditation. Although thesis studies on the concept of accreditation have been conducted across a range of disciplines, including distance education, business administration, educational programs, landscape architecture, visual arts, sports management, English language education, educational administration and supervision, it is evident that the

discipline with the highest number of studies is educational sciences. A total of eight theses were evaluated, four in the field of educational programs and teaching and four in the field of educational administration and supervision. This figure represents approximately half of the total number of theses created within the framework of the accreditation concept. The content descriptions related to the content of the studies are provided in Figure 2.

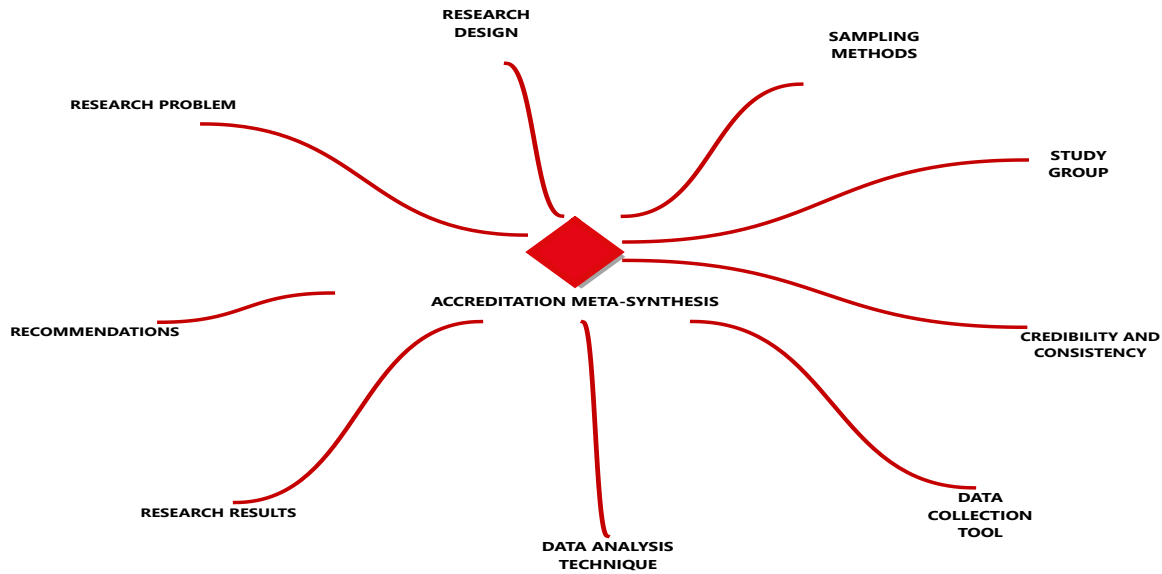


Figure 2. Code Co-Creation Model Related to the Content of the Studies

Figure 2 illustrates the research problem, research design, sampling method, study group, credibility and consistency, data collection tool, data analysis technique, research results and recommendations. The code-subcode section model, which serves to illustrate the research problems addressed in the thesis, is presented in Figure 3.

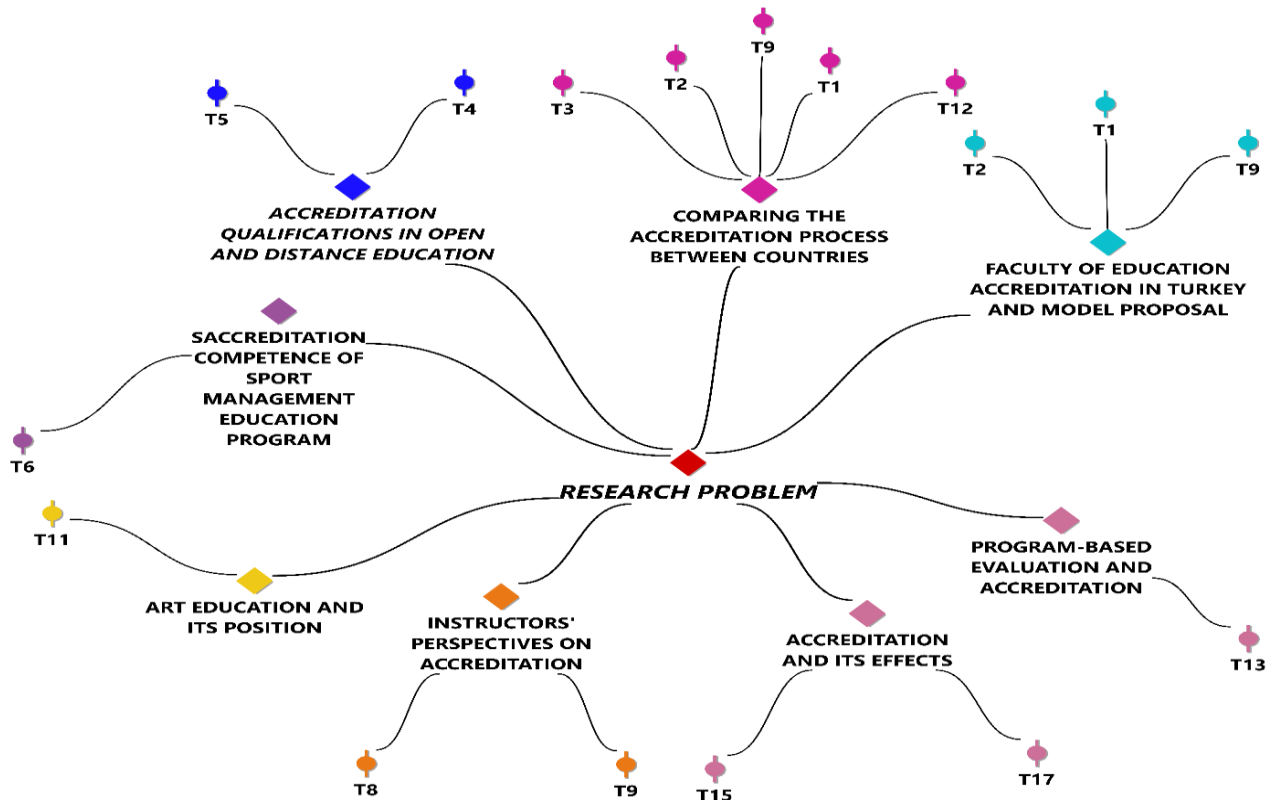


Figure 3. Code-Subcode Section Model for the Research Problems of the Theses

Upon closer examination of Figure 3, it becomes evident that the theses on the accreditation process encompass a range of codes pertaining to the theme of research problems. These include, but are not limited to, the following: comparison of accreditation processes between countries, accreditation of education faculties in Turkey and model proposal, art education and

its position, program-based evaluation and accreditation, accreditation and its effects, instructors' perspectives on accreditation, accreditation competence of sport management education program, and accreditation competencies in open and distance education. The most prevalent research topic identified was the comparative analysis of accreditation processes across countries, encompassing studies (T1-T2-T3-T9-T12). Secondly, the accreditation of education faculties in Turkey and the model proposal were subjected to detailed examination, and the research problems were structured as (T1-T2-T9). It was established that the code of accreditation competencies in the field of open and distance education (T4-T5) exhibited deficiencies in terms of research methodology. In the code of instructors' perspectives on accreditation (T8-T9), the issues they encounter are examined in detail. The accreditation process and its associated effects (T15-T17) were identified as problematic. The objective of this study was to ascertain the answer to the research question set forth in the program-based evaluation and accreditation code (T13). However, the question was addressed as a research question pertaining to art education and its position code (T11). In order to ascertain the adequacy of the accreditation code for the sport management education program (T6), an attempt was made to respond to the research question. The code-subcode section model, which pertains to the research design, is illustrated in Figure 4.

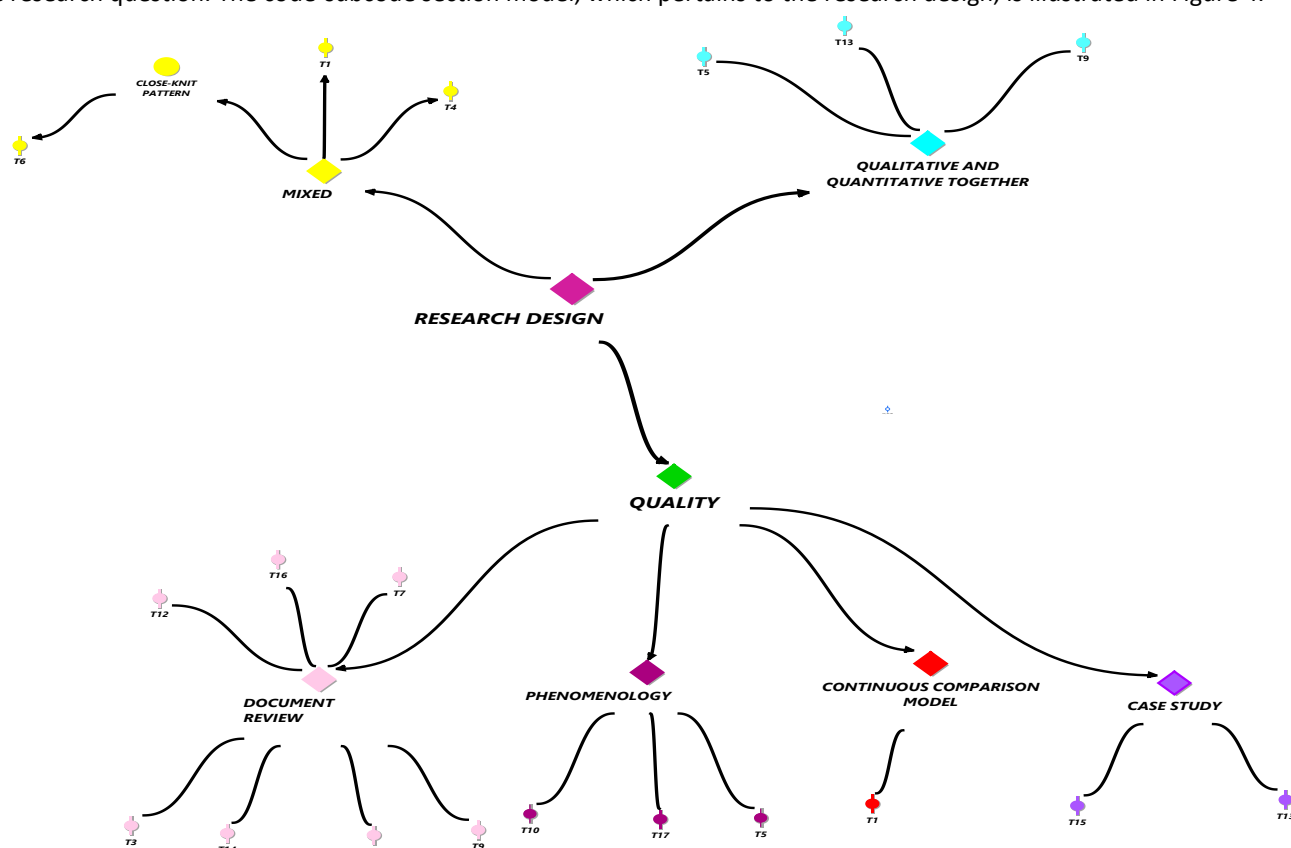


Figure 4. Code-Subcode Section Model Related to the Research Design of the Theses

As illustrated in Figure 4, the research design theme, situated within the overarching framework of an accreditation meta-synthesis, encompasses a diverse array of qualitative, mixed, qualitative, and quantitative codes. In the studies examined, it is evident that the convergent design is employed exclusively in the context of mixed methods (T6), whereas in other studies, the term "mixed" is used without specifying the particular designs of mixed methods (T1-T4). Furthermore, some theses indicate the use of both qualitative and quantitative methods without employing a mixed method approach. This is observed in studies that appear to be conducted in isolation from one another (T5-T13-T9).

In the case of theses in which the qualitative method was preferred, the document analysis design was identified as the most preferred approach (T9-T2-T14-T12-T16-T7). The second most utilized design is phenomenological studies (T5-T10-T17). In qualitative research, the case study (T13-T15) and constant comparison model (T1) were selected as the preferred methods. The code-subcode section model pertaining to the sampling methods utilized in the analyzed theses is illustrated in Figure 5.

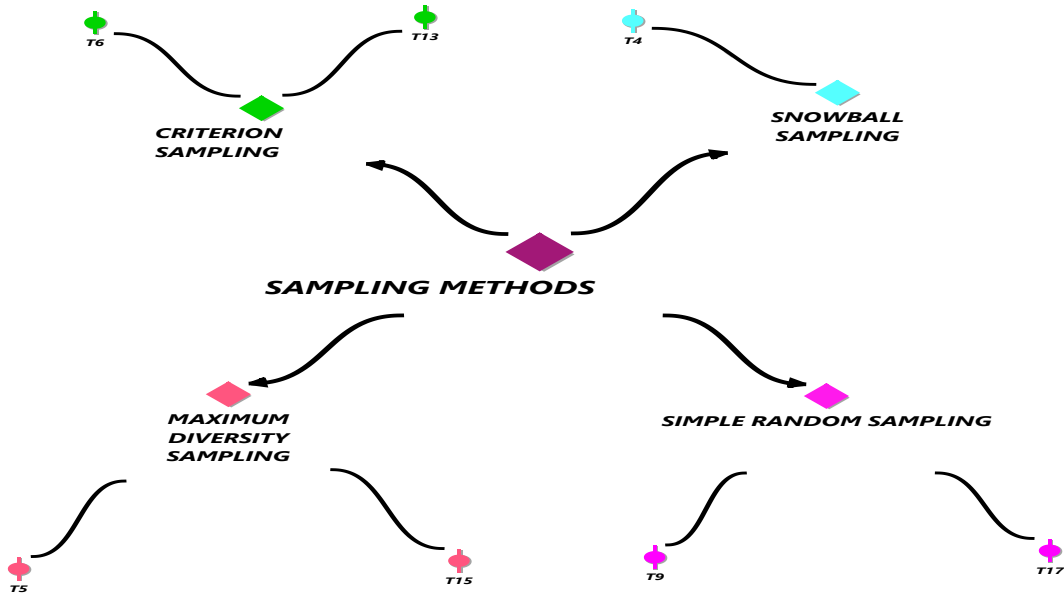


Figure 5. Code-Subcode Section Model Related to Sampling Methods of Theses

As illustrated in Figure 5, the sampling methods theme encompasses four distinct sampling techniques: criteria sampling, maximum variation sampling, snowball sampling, and simple random sampling. The sampling methods identified in the theses included criterion sampling (T6-T13), maximum variation sampling (T5-T15), snowball sampling (T4), and simple random sampling (T9-T17). The remaining theses did not explicitly describe the sampling methods used. The code-subcode section model pertaining to the study group of the analyzed theses is illustrated in Figure 6.

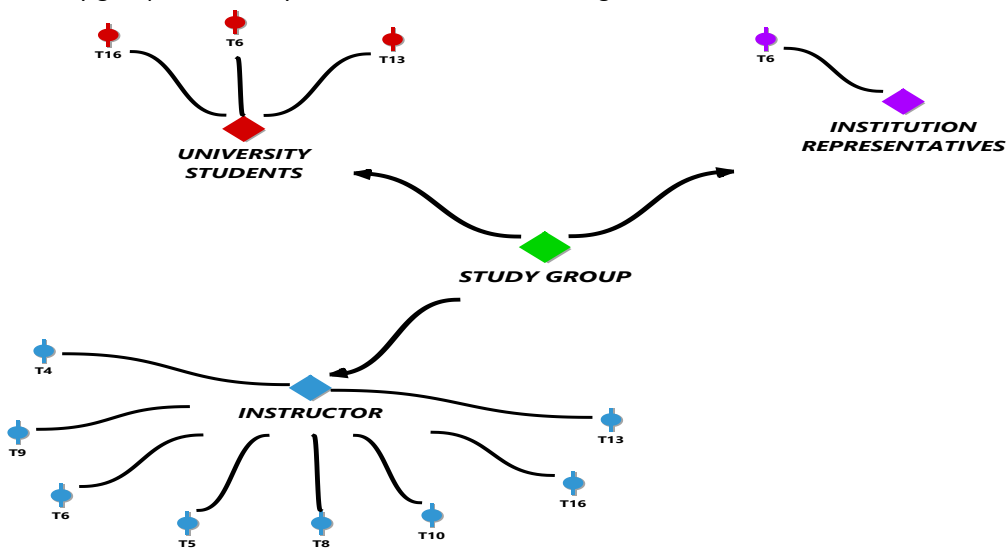


Figure 6. Code-Subcode Section Model Related to the Study Group of the Theses

Upon analysis of Figure 6, it becomes evident that the most favored study group was that comprising the academic staff (T4-T6-T5-T8-T10-T16-T9-T13). The instructors are followed by the students, in the following sequence: (T16, T6, T13). Furthermore, in thesis T6, institutional representatives were incorporated into the study group. The code-subcode section model pertaining to credibility and consistency is illustrated in Figure 7.

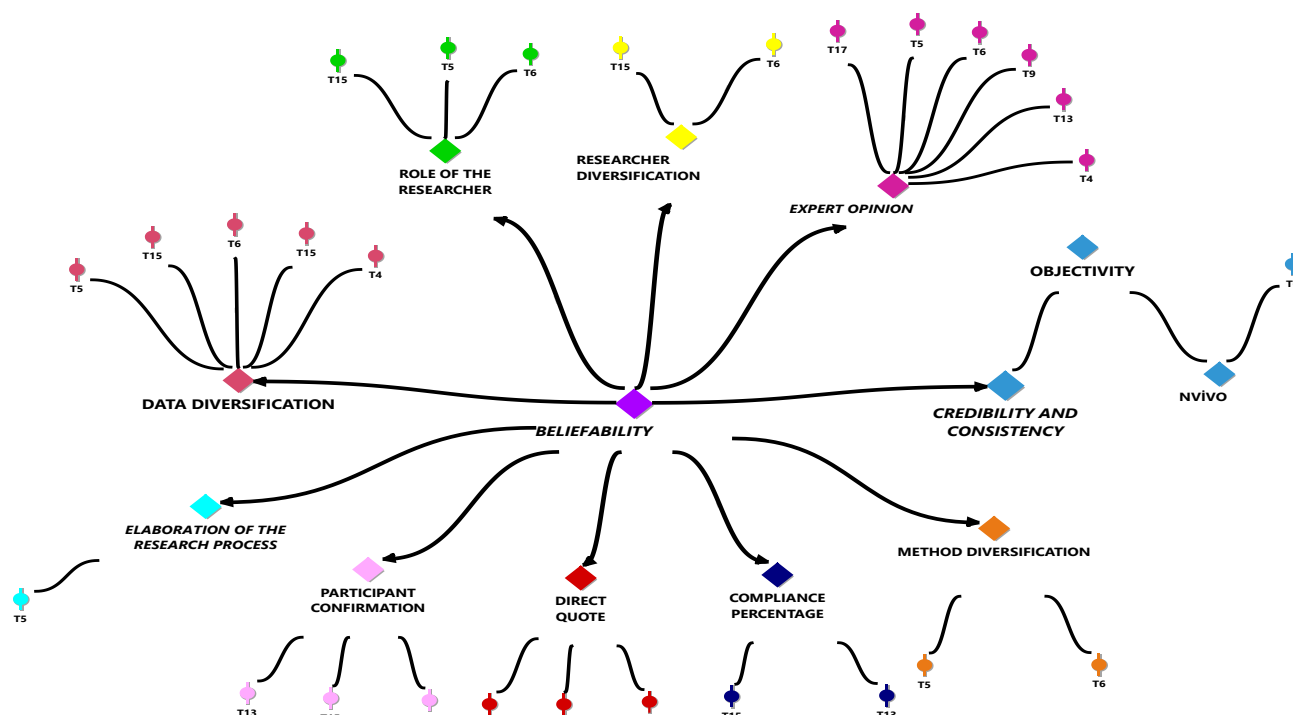


Figure 7. Code-Subcode Section Model of the Theses Related to Credibility and Consistency

As illustrated in Figure 7, the credibility and consistency theme encompasses objectivity and credibility codes. It was asserted that objectivity was incorporated solely into the thesis designated as T6 and that the qualitative analysis software (NVIVO) was employed to guarantee objectivity. In the context of credibility, expert opinion (T17, T5, T6, T9, T13, T4), researcher variation (T15, T6), the role of the researcher (T15, T5, T6), and data variation (T5, T15, T6) were identified as key considerations. 4) The subcodes include an elaboration of the research process (T5), confirmation of participants (T13, T15, T5), direct quotations (T15, T17, T13), the percentage of agreement (T15, T13), and variations in methodology (T5, T6). The code-subcode section model for the data collection tool is illustrated in Figure 8.

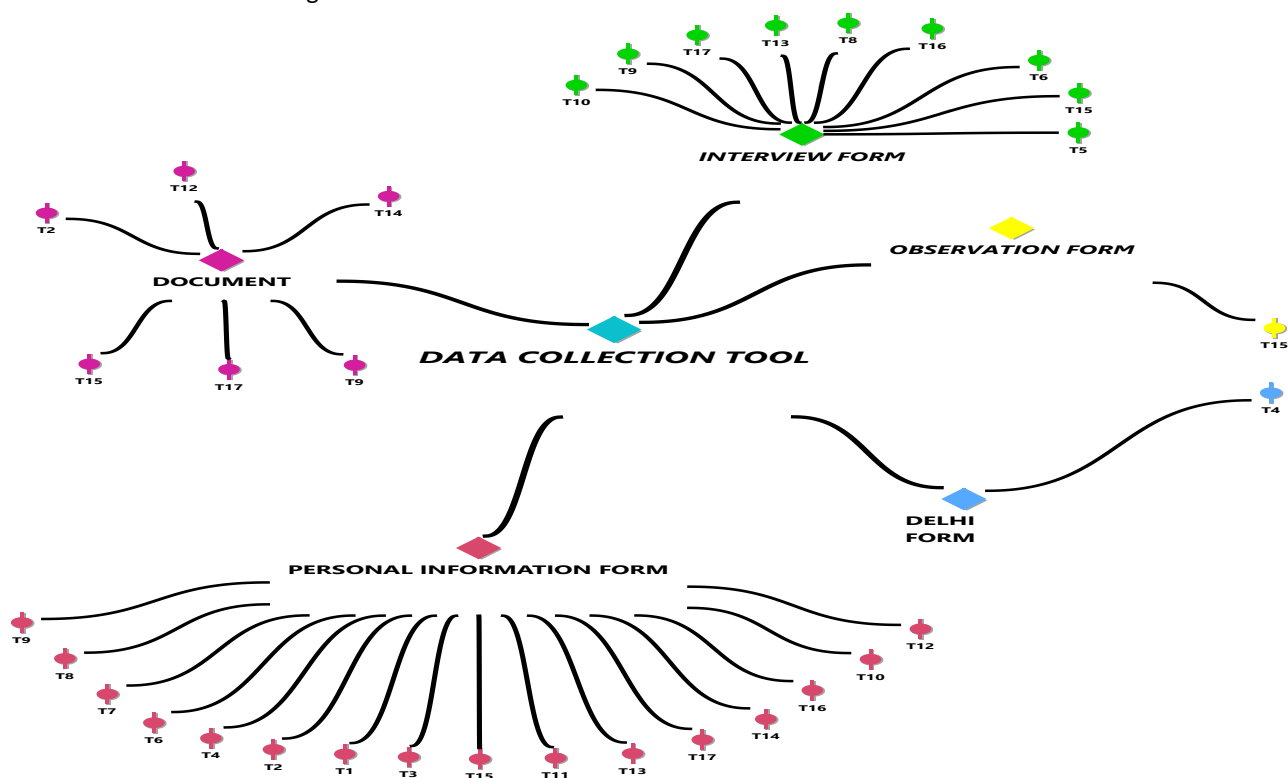


Figure 8. Code-Subcode Section Model of the Theses Related to the Data Collection Tool

As illustrated in Figure 8, personal information forms were utilized as data collection instruments in the entirety of the theses. (T6-T7-T8-T9-T4-T1-T2-T3-T15-T11-T13-T17-T14-T16-T10-T12). Following the administration of the personal information forms, the interview forms were utilized (T10-T17-T9-T9-T13-T13-T8-T16-T6-T15-T5). Document analysis was another preferred method for data collection (T12-T2-T14-T14-T15-T17-T9). In some instances, the observation form (T15), focus group interview form, and Delphi form were selected as the preferred methodology (T4). The code-subcode section model pertaining to the data analysis technique is illustrated in Figure 9.

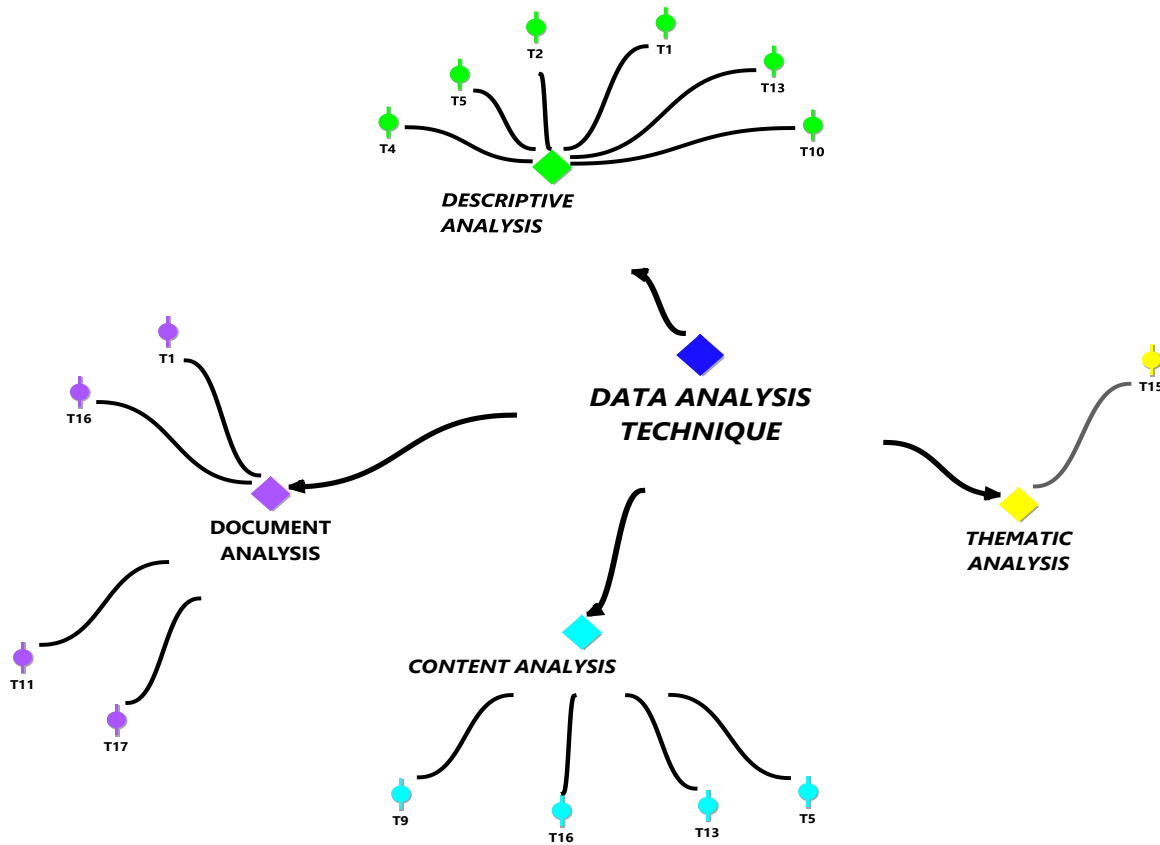


Figure 9. Code-Subcode Section Model of the Theses Related to Data Analysis Technique

As illustrated in Figure 9, descriptive analysis is the most preferred data analysis technique of the program, as evidenced by the responses from T4, T5, T2, T1, T13, and T10. Subsequently, a descriptive analysis is conducted, which is then followed by a content analysis (T9-T16-T13-T5) and a document analysis (T17-T11-T16-T1). A review of the thesis revealed that only the T15 thesis underwent a thematic analysis. The code-subcode section model pertaining to the findings of the analyzed theses is illustrated in Figure 10.

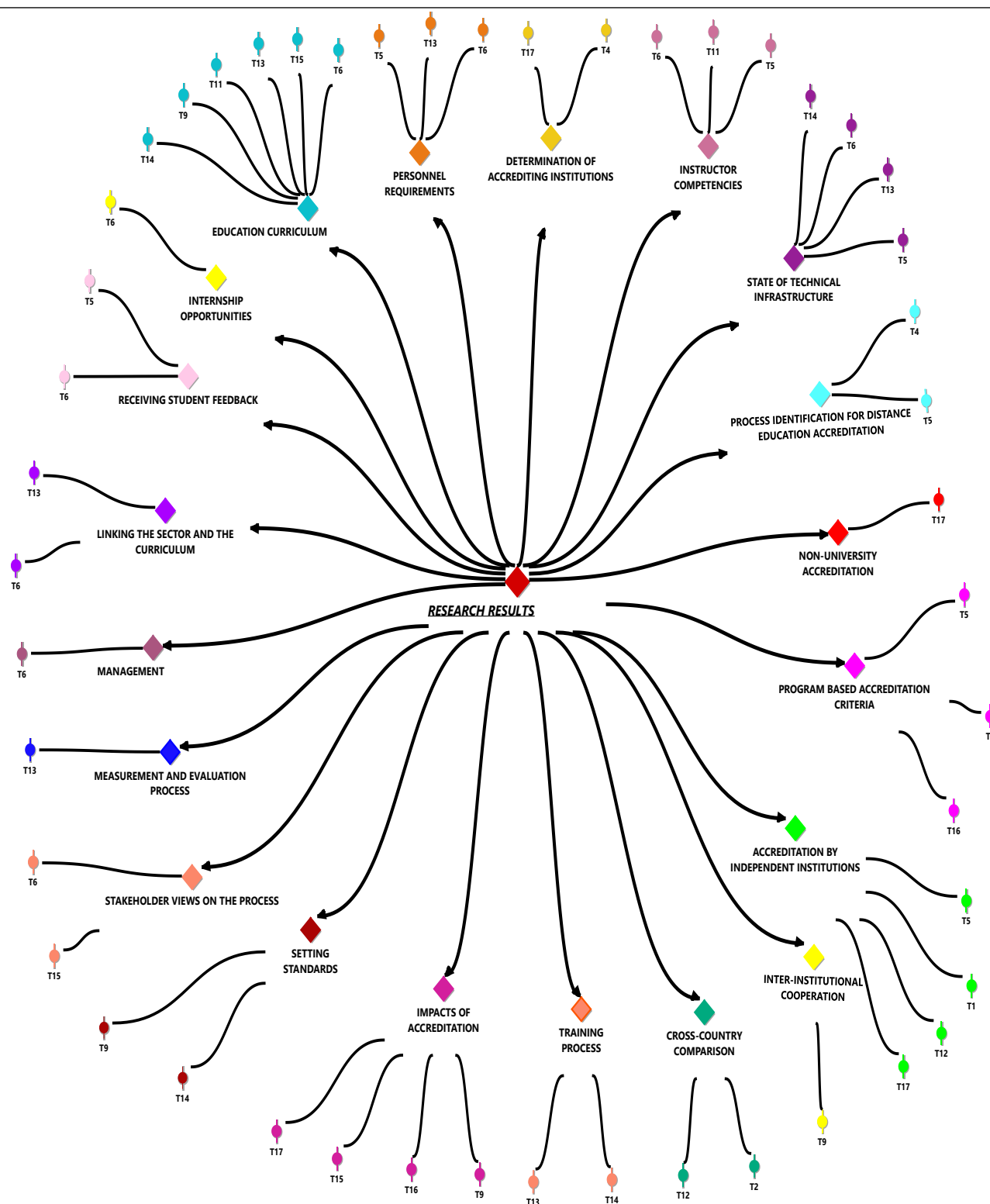


Figure 10.Code-Subcode Section Model Related to the Results of the Theses

Upon analysis of Figure 10, it becomes evident that the research results are related to several key themes, including instructor competencies (T6-T11 and T5), the determination of institutions to be accredited (T17-T4), personnel needs (T5-T13-T6), the training curriculum (T6-T15-T13-T11-T9-T14), internship opportunities (T6), and the receipt of Student feedback (T5-T6), alignment of the curriculum with industry standards (T13-T6), management (T6), assessment and evaluation processes (T13), stakeholder perspectives on the process (T6-T15), and the establishment of benchmarks (T9-T14). The codes encompass the impact of accreditation (T17-T15-T16-T9), the educational process (T13-T14), and cross- The following themes were identified in the analysis of Figure 10: country comparison (T12-T2), inter-institutional cooperation (T9), accreditation by independent institutions (T5-T1-T12-T17), accreditation criteria on a program basis (T5-T3-T16), external accreditation (T17), determining the process for distance education accreditation (T4-T5), and the state of technical infrastructure (T14-T6-T13-T5). The proposed code-subcode section model for the analyzed theses is illustrated in Figure 11.

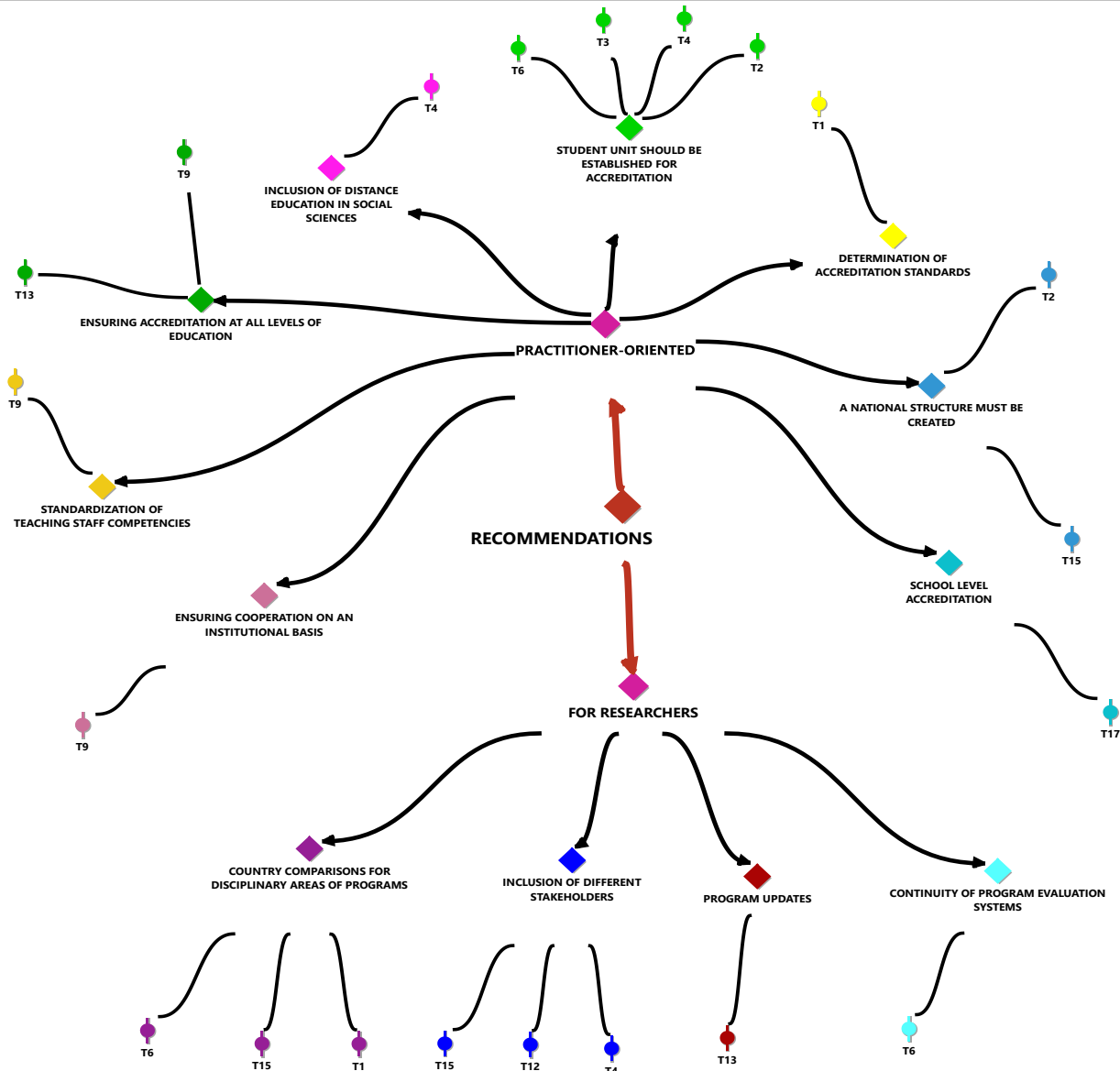


Figure 11.Cod Suggestions for Theses Code-Subcode Chapter Model

Upon examination of Figure 11, it becomes evident that the codes encompass recommendations for both practitioners and researchers. A review of the code of recommendations for practitioners reveals the presence of sub-codes, including the establishment of a student unit for accreditation (T2-T3-T4-T6), the integration of distance education in social sciences (T4), and the assurance of accreditation at all levels. The levels of education (T9-T13), standardization of instructor competencies (T9), assurance of institutional cooperation (T9), accreditation at the school level (T17), establishment of a national structure (T2-T15), and determination of accreditation standards. The theme of suggestions for researchers comprises the following sub-codes: country comparisons (T1-T6-T15), including different stakeholders (T15-T12-T4), program updates (T13), and continuity of program evaluation systems (T6).

DISCUSSION, CONCLUSION AND SUGGESTIONS

This study undertook a comprehensive examination of the problem situations, results, and recommendations identified in qualitative studies on the accreditation process as presented in the relevant literature. Upon examination of the research results, it became evident that numerous disciplines have conducted studies on the concept of accreditation. A review of the literature reveals that the field of study with the greatest emphasis on the concept of accreditation is that of educational sciences. It was posited that accreditation studies are instrumental in establishing and safeguarding the quality standards of educational institutions. The aforementioned standards encompass curriculum, faculty qualifications, student assessment, student support services, physical resources, and administration. From this perspective, it can be argued that the fields of accreditation and educational sciences are essentially synonymous, and that accreditation in the field of education is of critical importance (Kavurucu, 2024).

Given that the primary objective of accreditation is to enhance the quality of education and that it is based on educational curricula, student competencies, the learning-teaching process, and instructor competencies, it is reasonable to conclude that the field that bears the greatest responsibility is educational sciences.

Upon examination of the objectives set forth in the theses, it became evident that the majority of them sought to examine and contrast the processes of accreditation across different countries. Blanco (2019) and Nguyen, Marshall, and Evers (2021) underscore the significance of conducting cross-country comparisons of accreditation systems to identify shared characteristics, distinctive features, and recurrent themes within the accreditation process. Duarte & Vardasca (2023) and Reddy, Sharma & Gupta (2024) have asserted that a comparative analysis of accreditation processes across countries offers a unique opportunity to enhance the efficiency of accreditation procedures, identify exemplary practices, and provide guidance to policymakers and educational institutions in the pursuit of improved accreditation processes.

The findings indicated that document analysis was predominantly employed in accreditation studies, with qualitative methods being the preferred approach. Schomaker (2015) posited that document analysis enables a comparative and systematic examination of the evolution of accreditation standards across diverse countries and fields over time, as well as an investigation of pertinent country and institutional policies and reports on this subject. In a recent publication, Penman, Raymond, Kumar, Liang, Sundar, Kumar, Liang, Sundar, and Thomas (2023) posited that document analysis provides a valuable means of evaluating the relative merits and shortcomings of the content of training programs and their alignment with national qualification standards. In their 2021 study, Azzam, Puvirajah, Girard, and Grymonpre asserted that document analysis enables the tracing of the objectives, contents, and historical development of accreditation systems across diverse sectors. Wike and Cahyasari (2019) indicate that document analysis is frequently the preferred method for determining how academic staff quality is assessed and for identifying the commonalities and relationships between the methods used to evaluate staff quality.

A review of the literature reveals that the majority of accreditation studies are conducted on academics. As reported by Salto (2022) and Kadir, Arshad, and Johari (2016), academic staff are preferred due to their demonstrated expertise in curriculum design, teaching methodologies, self-assessment processes, and learning outcomes.

In the qualitative studies on the accreditation process, it was determined that expert opinion and data triangulation were the most frequently utilized methods for establishing credibility. While expert review is regarded as a crucial step to guarantee the objectivity of the data and the integrity of the analysis and research process, Creswell and Poth (2018) asserted that it is a vital element for the validity of the study. The rationale behind the preference for data triangulation in qualitative research lies in its capacity to enhance the credibility of the findings. As posited by Patton (2014), this approach can be regarded as a crucial step in augmenting the validity of the research, whereby the strengths of one approach can offset the shortcomings of another.

The findings of the analysis of the theses indicated that the educational curriculum was the most effective. Upon examination of the literature, Frank et al. (2020) underscored the pivotal role of accreditations in shaping the quality of educational programs, upholding the standards of such programs, aligning the curriculum with societal needs, and enhancing learning environments. Furthermore, Kayingo, Gordes, Kulo, and Cawley (2022) asserted that accreditation is an efficacious method for guaranteeing that educational programs adhere to the requisite standards, maintaining pedagogical integrity and uniformity across the program, and providing guidance on academic policies and practices.

The theses posit that accreditation should be conducted by autonomous institutions. Gutiérrez, Rodríguez, Torres-Pereda, and Reyes-Morales (2024) underscored the significance of accreditation by independent institutions in their study of the Mexican health sector. The discussion centered on the potential influence of administrations on the accreditation process, leading to the observation that accreditation is not perceived as a quality process but rather as a mechanism for funding. The necessity of independent, politically and otherwise unencumbered accrediting bodies to ensure the quality of research was a point of emphasis.

A noteworthy finding in the theses was the correlation between technical infrastructure and accreditation. Joseph (2021) underscored the necessity of enhancing infrastructure and personnel facilities to ensure the provision of a quality educational service. In their study, Holt, Miller, Byrne, and Day (2019) emphasized that accreditation led to the establishment of learning and clinical environments in medical education that were not available in 2010. These infrastructure improvements had a positive effect on the quality of education in subsequent studies. In their study on the accreditation of medical education in Pakistan, Khan, Sethi, Wajid, and Yasmeen (2020) posited that, in order to enhance the quality of medical education, it is imperative that the theoretical structure be strengthened and that accreditation organizations implement reforms.

The most prevalent recommendation for practitioners is the establishment of a student unit for accreditation. Spowart & Turner (2022) posit that students should be active participants in the development of their curriculum, serve as collaborators in the accreditation process, and engage in accreditation evaluations. They contend that this approach will enhance transparency and positively impact the quality of education in the long term.

Kafaji (2020) posited that accreditation awareness programs, in which students are actively involved, can enhance their potential to achieve accreditation. Choa, Arfeen, Chan, & Rashid (2022) posited that administrators ought to consider student perspectives when implementing and enhancing the accreditation program.

Among the recommendations put forth for researchers, the most prevalent include country comparisons for the disciplinary areas of the programs and the inclusion of various stakeholders. Gevorkyan & Nikitin (2021) conducted a comparative analysis of

accreditation processes and criteria in various countries, including Canada, Hong Kong, Australia, and the USA. Their study revealed that the presence of departmental accreditation on a program basis varies across nations, with some countries implementing such accreditation systems and others not. Additionally, the analysis identified instances where countries exhibit internal variations in their accreditation frameworks. The study concludes with the observation that there is currently no uniformity in the accreditation standards employed by these countries.

In their study, Allen, Baalawi, Al Shoaibi, Gomma, and Rock (2022) compared the accreditation of medical education in northern Europe and the Arab Emirates, examining the accreditation criteria on a program-by-program basis. The study under scrutiny posits that the comparison in question can be regarded as a paradigm for the purpose of mitigating discrepancies between nations and compensating for the shortcomings of each on an individual basis. It has been documented that the accreditation standards established by the US medical education committee are applicable on an international scale.

In view of the results of this study, a number of recommendations have been formulated to guide practitioners and researchers in their endeavors. Firstly, the establishment of student units is recommended to ensure the active involvement of students in the accreditation process. This is necessary to carry out the accreditation process in a more effective and inclusive manner. These units have the potential to enhance student involvement in the development and evaluation of curricula, thereby ensuring a more transparent process and contributing to the enhancement of educational quality. The necessity of accreditation by autonomous and independent institutions is also emphasized, and the recommendation is made for the establishment of an objective and transparent audit system. The enhancement of technical and physical infrastructure within educational institutions is regarded as imperative for ensuring optimal learning environments, particularly in applied fields. Furthermore, curricula must be structured to align with contemporary social and sectoral requirements. It is imperative to elucidate learning outcomes and ensure that instructors undergo ongoing training on subjects such as accreditation, quality management, and self-assessment.

For researchers, the necessity of increasing discipline-based and cross-country comparative studies has been asserted. Such studies can contribute to the development of a model for good practices by comparing accreditation processes in different contexts. In subsequent research, incorporating a diverse range of stakeholders, including students, administrators, quality assurance units, and graduates, will enhance the comprehensiveness and validity of the data collected. Furthermore, longitudinal or mixed-methods studies will provide significant contributions in terms of evaluating the effects of accreditation over time and integrating quantitative and qualitative dimensions. However, the execution of more comprehensive and multifaceted meta-synthesis studies is recommended, incorporating not only theses or articles, but also policy documents, legislation, and accreditation reports. Research to be developed in this context will contribute to the production of theoretical knowledge and form the basis for the formulation of practical policy recommendations.

Declaration of Conflicting Interests

The author declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author received no financial support for the research, authorship, and/or publication of this article.

Statements of publication ethics

I hereby declare that the study has not unethical issues and that research and publication ethics have been observed carefully.

Researchers' contribution rate

The researcher was responsible for all stages of the research process, including the development of the theoretical framework, determination of the method and design, data collection, data analysis, interpretation of findings, and writing of the article. All phases of the study were conducted independently, and the final version of the article was approved by the researcher.

Ethics Committee Approval Information

The present study was conducted using a combination of documentary review (document analysis) and meta-synthesis methods. The study examined only publicly available theses. No data was collected directly from human or animal participants. Consequently, ethics committee approval is not a prerequisite. All data utilized in the study were obtained from open-access sources, and ethical principles were adhered to throughout the research process.

REFERENCES

- Allen, S. K., Baalawi, Z. S., Al Shoaibi, A., Gomma, H. W., & Rock, J. A. (2022). Applying North American medical education accreditation standards internationally in the United Arab Emirates. *Medical Education Online*, 27(1), 2057790. <https://doi.org/10.1080/10872981.2022.2057790>
- Asiyai, R. I. (2015). Improving quality higher education in Nigeria: The roles of stakeholders. *International Journal of Higher Education*, 4(1), 61–70. <https://doi.org/10.5430/ijhe.v4n1p61>
- Azzam, M., Puvirajah, A., Girard, M. A., & Grymonpre, R. E. (2021). Interprofessional education-relevant accreditation standards in Canada: A comparative document analysis. *Human Resources for Health*, 19(1), 66. <https://doi.org/10.1186/s12960-021-00622-w>
- Blanco, G. (2019). U.S. accreditation as international accreditation: Casting U.S. practices as global standards and "best practices" for quality assurance. <https://doi.org/10.3102/1439065>
- Choa, G., Arfeen, Z., Chan, S. C. C., & Rashid, M. A. (2022). Understanding impacts of accreditation on medical teachers and students: A systematic review and meta-ethnography. *Medical Teacher*, 44(1), 63–70. <https://doi.org/10.1080/0142159X.2021.1912522>
- Creswell, J. W., & Poth, C. N. (2018). *Qualitative inquiry and research design: Choosing among five approaches*. Sage Publications.
- Duarte, N., & Vardasca, R. (2023). Literature review of accreditation systems in higher education. *Education Sciences*, 13(6), 582. <https://doi.org/10.3390/educsci13060582>
- Eaton, J. S. (2015). *An overview of US accreditation* (Revised November 2015). Council for Higher Education Accreditation.
- Ebekozien, A., & Aigbavboa, C. (2023). Evaluation of built environment programmes accreditation in the 21st century education system in Nigeria: Stakeholders' perspective. *International Journal of Building Pathology and Adaptation*, 41(6), 102–118. <https://doi.org/10.1108/IJBPA-03-2023-0102>
- Fernandes, J. O., & Singh, B. (2022). Accreditation and ranking of higher education institutions (HEIs): Review, observations, and recommendations for the Indian higher education system. *The TQM Journal*, 34(5), 1013–1038. <https://doi.org/10.1108/TQM-03-2021-0101>
- Finfeld-Connett, D. (2010). Generalizability and transferability of meta-synthesis research findings. *Journal of Advanced Nursing*, 66(2), 246–254. <https://doi.org/10.1111/j.1365-2648.2009.05283.x>
- Frank, J. R., Taber, S., van Zanten, M., Scheele, F., Blouin, D., & International Health Professions Accreditation Outcomes Consortium. (2020). The role of accreditation in 21st century health professions education: Report of an international consensus group. *BMC Medical Education*, 20, 1–9. <https://doi.org/10.1186/s12909-020-02284-7>
- Gevorkyan, E. N., & Nikitin, I. V. (2021). Accreditation and teacher education: A cross-national case study. *European Proceedings of Educational Sciences*, 5, 101–111. <https://doi.org/10.15405/epes.22043.10>
- Gutiérrez, J. P., Rodriguez, M. A., Torres-Pereda, P., & Reyes-Morales, H. (2024). Hospital accreditation in Mexico fails to improve the quality of healthcare: Lessons from an impact evaluation. *Frontiers in Public Health*, 12, 1–8. <https://doi.org/10.3389/fpubh.2024.1386667>
- Holt, K. D., Miller, R. S., Byrne, L. M., & Day, S. H. (2019). The positive effects of accreditation on graduate medical education programs in Singapore. *Journal of Graduate Medical Education*, 213–217. <https://doi.org/10.4300/JGME-D-18-00879.1>
- Joseph, S. (2021). Does accreditation symbolise quality in public healthcare delivery? An investigation of hospitals in Kerala. *Journal of Health Management*, 23(4), 715–729. <https://doi.org/10.1177/0972063421993365>
- Kadir, K. A., Arshad, D. A., & Johari, J. (2016). Resistance to change of academics towards accreditation. *Journal of Business and Social Review in Emerging Economies*, 2(2), 127–134. <https://doi.org/10.26710/jbseer.v2i2.31>
- Kafaji, M. (2020). The perceived benefits of accreditation on students' performance: The case of private business schools. *Industry and Higher Education*, 34(6), 421–428. <https://doi.org/10.1177/0950422220968085>
- Kavurucu, O. (2024). Evaluation of accreditation studies at primary and secondary education level in Turkey and Europe. *Journal of Social, Humanities and Administrative Sciences (JOSHAS)*, 9(65), 3069–3074.
- Kayingo, G., Gordes, K. L., Kulo, V., & Cawley, J. F. (2022). Accreditation in the health professions: Implications for physician assistant education. *The Journal of Physician Assistant Education*, 33(4), 318–324. <https://doi.org/10.1097/JPA.0000000000000436>
- Khan, A. W., Sethi, A., Wajid, G., & Yasmeen, R. (2020). Challenges towards quality assurance of basic medical education in Pakistan. *Pakistan Journal of Medical Sciences*, 36(2), 4. <https://doi.org/10.12669/pjms.36.2.4>
- Kundu, G. K., & Majumdar, J. P. (2020). The learning and teaching area of AACSB standards: A process model framework. *Business Process Management Journal*, 26(6), 1379–1399. <https://doi.org/10.1108/BPMJ-09-2019-0367>
- Lagrosen, S. O. (2017). Quality through accreditation. *International Journal of Quality and Service Sciences*, 9(3/4), 469–483. <https://doi.org/10.1108/IJQSS-01-2017-0023>
- López Segrera, F., Sanyal, B. C., & Tres, J. (2007). An overview of regional perspectives on accreditation in today's world. *Higher Education in the World* 2007, 1, 179–187.
- Mati, Y. (2018). Input resources indicators in use for accreditation purpose of higher education institutions. *Performance Measurement and Metrics*, 19(3), 176–185. <https://doi.org/10.1108/PMM-01-2018-0034>
- Nguyen, C. H., Marshall, S. J., & Evers, C. W. (2021). Higher education quality assurance and accreditation implementation in several countries across the world and lessons learned for Vietnam. *Vietnam Journal of Education*, 11–17.
- Nicklin, W., Engel, C., & Stewart, J. (2021). Accreditation in 2030. *International Journal for Quality in Health Care*, 33(1), 1–5. <https://doi.org/10.1093/intqhc/mzaa15>
- Patton, M. Q. (2002). *Qualitative research & evaluation methods*. Sage Publications.
- Patton, M. Q. (2014). *Qualitative research & evaluation methods: Integrating theory and practice*. Sage Publications.

- Penman, M., Raymond, J., Kumar, A., Liang, R. Y., Sundar, K., & Thomas, Y. (2023). Allied health professions accreditation standards for work integrated learning: A document analysis. *International Journal of Environmental Research and Public Health*, 20(15), 64-78. <https://doi.org/10.3390/ijerph20156478>
- Reddy, J. S., Sharma, R., & Gupta, N. (2024). The accreditation paradigm: A comparative analysis of accreditations for management programmes. *International Journal of Educational Management*, 38(1), 73–95. <https://doi.org/10.1108/IJEM-07-2022-0302>
- Salto, D. J. (2022). Do academic disciplines matter? An analysis of organizational responses to the accreditation of graduate programs by field of study and sector. *Higher Education*, 84(3), 569–587. <https://doi.org/10.1007/s10734-022-00850-4>
- Sánchez-Chaparro, T., Remaud, B., Gómez-Frías, V., Duykaerts, C., & Jolly, A. M. (2022). Benefits and challenges of cross-border quality assurance in higher education: A case study in engineering education in Europe. *Quality in Higher Education*, 28(3), 308–325. <https://doi.org/10.1080/13538322.2022.2129702>
- Schomaker, R. (2015). Accreditation and quality assurance in the Egyptian higher education system. *Quality Assurance in Education*, 23(2), 149–165. <https://doi.org/10.1108/QAE-10-2013-0047>
- Spowart, L., & Turner, R. (2022). Institutional accreditation and the professionalisation of teaching in the HE sector. In *Higher education—New approaches to accreditation, digitalization, and globalization in the age of Covid 1–17*.
- Wike, W., & Cahyasari, E. (2019, August). The quality of higher education assessed from the academic staff performance: Literature review. In *Annual International Conference of Business and Public Administration (AICoBPA 2018)*, 253–256. Atlantis Press. <https://doi.org/10.2991/aicobpa-18.2019.44>



| Research Article / Araştırma Makalesi |

Investigation of Studies on Micro-credentials by Bibliometric Analysis Method

Mikro Yeterlikler ile İlgili Yapılmış Çalışmaların Bibliyometrik Analiz Yöntemiyle İncelenmesi

Akça Okan YÜKSEL¹

Keywords

1. Micro-credentials
2. Higher Education
3. Digital badge
4. Bibliometric analysis

Anahtar Kelimeler

1. Mikro yeterlikler
2. Yüksek öğretim
3. Dijital rozet
4. Bibliyometrik analiz

Received/Başvuru Tarihi

15.01.2025

Accepted / Kabul Tarihi

10.10.2025

Abstract

Purpose: The aim of the study is to reveal the trend of micro-credential studies conducted at the higher education level and to present a general framework on micro-credentials.

Design/Methodology/Approach: Bibliometric analysis method was used in the study. Web of Science (WOS) database, which is a widely used and reliable data source in literature searches, was preferred to search the articles. The search criteria were limited to articles with title, abstract and keywords. For this purpose, the Web of Science database was searched with the keywords 'micro-credentials' and 'micro credential'. 85 papers are identified and included in the analysis.

Findings: The keyword analysis revealed that micro-credentials have strong links with higher education, digital badges, employment and lifelong learning. At the same time, the most cited countries and researchers were mostly from countries such as the USA, Australia and Canada. The co-occurrence analysis of keywords unveils five clusters representing trends in micro-competencies and employability, learning and skills, distance learning, digitalization and technology, quality assurance and sustainability.

Highlights: In conclusion, micro-credentials are becoming increasingly important as a flexible and innovative learning model that responds to the needs of modern education systems and labor markets.

Öz

Çalışmanın amacı: Çalışmanın amacı, yükseköğretim düzeyinde mikro yeterlikler üzerine yapılmış çalışmaların eğilimini ortaya koymak ve mikro yeterliklere ilişkin genel bir çerçeve sunmaktır.

Materyal ve Yöntem: Çalışmada bibliyometrik analiz yöntemi kullanılmıştır. Makalelerin taranmasında literatür taramalarında yaygın olarak kullanılan ve güvenilir bir veri kaynağı olan Web of Science (WOS) veri tabanı tercih edilmiştir. Arama kriterleri başlık, özet ve anahtar kelimelerin yer aldığı makalelerle sınırlandırılmıştır. Bu amaçla Web of Science veri tabanında 'micro-credentials' ve 'micro credential' anahtar kelimeleri ile arama yapılmıştır. Arama sonucunda, 85 makale tespit edilmiş ve analize dahil edilmiştir.

Bulgular: Yapılan analizler sonucunda, anahtar kelime analizi mikro yeterlikler kavramının yüksek öğrenim, dijital rozetler, istihdam ve yaşam boyu öğrenme ile güçlü bağlantıları olduğunu ortaya koymuştur. Aynı zamanda, en çok atıfta bulunulan ülkeler ve araştırmacılar çoğunlukla ABD, Avustralya ve Kanada gibi ülkelerden olmuştur. Anahtar kelimelerin birlikte ortaya çıkma analizi, mikro yeterlilikler ve istihdam edilebilirlik, öğrenme ve beceriler, uzaktan eğitim, dijitalleşme ve teknoloji, kalite güvencesi ve sürdürülebilirlik alanlarındaki eğilimleri temsil eden beş kümeyi ortaya çıkarmıştır.

Önemli Vurgular: Sonuç olarak, mikro yeterlikler modern eğitim sistemlerinin ve işgücü piyasalarının ihtiyaçlarına cevap veren esnek ve yenilikçi bir öğrenme modeli olarak giderek daha önemli hale gelmektedir.

¹ Corresponded Author, Middle East Technical University, Information System Coordination Office, Ankara, Türkiye; <https://orcid.org/0000-0002-5430-0821>

INTRODUCTION

In the business world, the skills and related employment gap is a major problem stemming from the mismatch between the skill sets of job seekers and the demands of employers. Digitalization and technological transformation are rapidly changing the skills demanded by businesses, but education and training systems are not responding quickly enough to this change (Tee et al., 20-24). In this context, micro-credentials stand out as a flexible and competency-based learning model that facilitates individuals to quickly adapt to changing expectations in the labor market. Micro-credentials is a competency-based learning model provided by higher education or business world that, upon completion, awards students with a digital badge (Kayyali, 2025). Micro-credentials is a systematic learning model provided by higher education or the business world that awards a digitally distributed badge upon completion (Kayyali, 2025). Micro-credentials refer to evidence of learning in which learning outcomes are assessed according to transparent standards. According to another definition, micro-credentials are formal qualification documents that enable individuals to acquire specific skills through short-term training modules and are awarded to individuals after assessment (Tan et al., 2024). It is seen that micro-credentials are demanded by employers, especially in areas such as digital skills, problem-solving competencies and security awareness (Rižnar, 2023). From the perspective of the business world, it is stated that micro-credentials are one of the important factors affecting employers' employee selection and employers prefer candidates with these competencies more (McGreal & Olcott, 2022). The competency-based, flexible and demand-oriented structure of micro-credentials is particularly effective in providing educational opportunities to disadvantaged groups. By offering an inclusive education model, it supports the personal development of individuals and promotes lifelong learning (Murgatroyd, 2022). Ward et al. (2023) emphasized in their study that micro-credentials offer great potential in terms of gaining new competencies in the business world and facilitating career transitions. Micro-credentials are growing in relevance in closing the skills gap in the business world. Considered as a valuable tool to fit the shifting dynamics of the labor market, boosts employability of individuals, and encourage lifelong learning are micro-credentials. In this framework, it becomes a vital need for policy makers, educational institutions, and the corporate sector to assist this new learning model more successfully. qualified workforce is one of the essential elements that companies should consider when planning (Şerbetçi, 2003). According to Eskandaricharati (2013), high employee motivation and business commitment are essential for businesses to continue their existence. Organizational commitment means a person's emotional responses to the characteristics of the organization they work for. Employees' organizational commitment is significant for business success (Kara, 2019; Majid & Afifa, 2017). According to Steers (1981), organizational commitment includes individuals' desire to take an active role in the success of their organization and the achievement of its goals beyond passive loyalty to the organization. Therefore, organizational commitment is a broad concept that includes employees' voluntary participation in activities that will move the company forward, and it means more than just the emotional bond established with the company. According to Uygur (2009), employees must believe in the organization's goals for organizational commitment to occur. This way, the employee strongly desires to work to benefit the organization. Organizational commitment is determined by the employees' relationship with the organization and is a psychological behavior that ensures their continued stay in the organization (Özdevecioğlu, 2003). Demir and Öztürk (2011) defined organizational commitment as individuals giving more importance to the organization's interests than other factors. According to Reichers (1985), the desire to stay in the organization results from organizational commitment.

Digital badges are one of the ideas related to micro-credentials. Micro-credentials are structures that, via small, observable learning units, record individuals' competency in each skill or knowledge area. Usually supported by block chain or trusted verification systems, digital badges are symbols that offer visual and digital depiction of these micro-credentials. Micro-credentials, which digital evidence of learning, skills, and competencies (Kukkonen, 2021) are given in higher education as digital badges, by enabling the shared micro-credentials on digital platforms, digital badges help people make their competencies visible to companies, educational institutions, and professional networks (Kır & Bozkurt, 2022). Micro-credentials and digital badges taken together then enable more flexible, transparent learning environments.

Within the framework of educational technologies, micro-credentials stand out as a creative assessment tool that lets people acquire and validate skills and competencies by means of small, modular learning units (Orman, Simsek & Çakır, 2023). Reliable and flexible micro-credential implementation is made possible by digital learning platforms, AI-powered customized education systems, and block chain-based validation systems (McGreal, 2023). This allows students to acquire the skills they need without depending on conventional long-term training courses and can show these improvements in professional networks by means of digital badges (Ahmat et al., 2021). Micro-credentials are increasingly important in individualized learning experiences and lifetime learning processes thanks to their adaptability, accessibility, and scalability given by educational technologies (Ali & Khan, 2023).

As a structure that supports flexible and modular educational approaches as an alternative to conventional diploma-oriented learning models, micro-credentials in higher education are progressively drawing attention in the educational sciences literature (Orman, Simsek & Çakır, 2023). By encouraging individualized learning experiences (Pirkkalainen et al., 2022), micro-credentials support learners' self-directed and needs-based skill acquisition, so making one of the most significant contributions they make to educational sciences. Simultaneously, the use of micro-credentials in higher education institutions integrates with digital learning environments and artificial intelligence-supported personalized education processes, so improving access to and adaptability of learning processes (Ali & Khan, 2023). In this regard, the use of micro-credentials in higher education is seen as an original model that supports lifetime learning for both instructional designers and policy makers (Ahmat et al., 2021).

The use of micro-credentials in literature is still very new and there are very few studies in which they are used for teaching purposes and their results are shared. The aim of the study is to reveal the trend of micro-credential studies conducted at the higher education level and to present a general framework on micro-credentials. Thus, as a result of the study, a holistic and inclusive perspective will be provided to the existing studies, and the results of the study will be able to guide the studies to be carried out in this field and will help to create a starting point for new studies. In this context, answers to the following research questions were sought.

- What is the distribution of publications within the scope of micro-credentials by year, journal, and research field?
- What is the distribution of publications made within the scope of micro-credentials according to sustainable development goals?
- What is the distribution of publications made within the scope of micro-credentials according to the main topics and concepts?
- What is the bibliographic matching of the publications made within the scope of micro-credentials and the authors of these publications?
- What is the distribution of publications made within the scope of micro-credentials according to keywords?
- How are the keywords clustered in publications made within the scope of micro-credentials?

METHOD/MATERIALS

The aim of this study is to bibliometrically analyze the micro-credentials studies conducted within the scope of educational sciences. Bibliometric analysis method was used in the study. This method includes mathematical and statistical methodology aimed at evaluating bibliometric data (Donthu et al., 2021). Since the aim of the study was to investigate research trends on micro-credentials in higher education, the unit of analysis was a single article focusing on micro-credentials. To search for the related articles, the Web of Science database, which is a widely used and reliable data source for literature searches, was chosen. The search criteria were restricted to articles with title, abstract and keywords. For this purpose, the Web of Science database was searched with the keywords “micro-credentials” and “micro credential”. Within the scope of the study, it was analyzed using the search strategy: (micro-credentials OR micro-credential OR microcredential) AND (educa* OR learn* OR teach*). These keywords formed the search protocol to identify literature on micro-credentials in higher education. The search was last completed on 01.02.2025.

When selecting the articles, inclusion and exclusion criteria were defined following the PRISMA guidelines for systematic literature reviews proposed by Page et al. (2021) (Table 1). A rigorous selection process was then carried out in three distinct stages: identification, screening and inclusion. This systematic approach ensured a comprehensive and rigorous selection of articles that met the research objectives.

Table 1. Inclusion and Exclusion Criteria

Inclusion Criteria
Published in an academic journal
Written in English
Available in full-text
Research paper in the higher education level
Exclusion Criteria
Proceedings, meeting abstracts, editorial materials

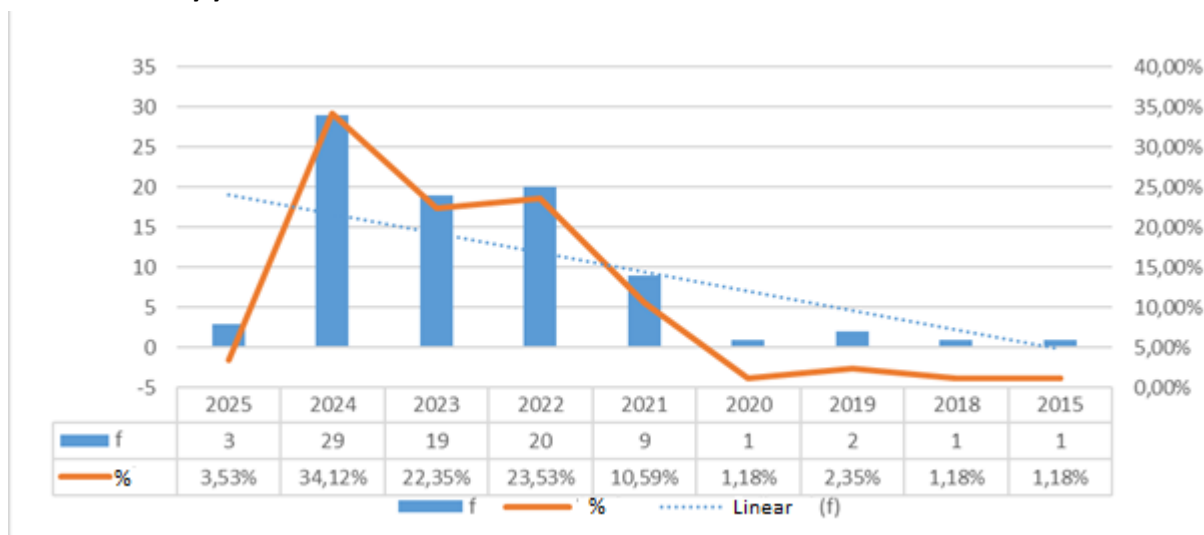
After the inclusion and exclusion criteria specified in the table, a total of 85 studies were identified and included in the analysis. Details of the query and selection processes are given in Table 2. WOSviewer, a well-known, reliable and valid software, was used to perform the analysis (Van Eck & Waltman, 2014).

Table 2. PRISMA Chart

Identification	Screening	Included
Articles identified of WOS	After selected title, abstract	After inclusion and exclusion
233	132	85

FINDINGS

Distribution of studies by year

**Figure 1. Distribution of Studies by Years**

When the distribution of academic studies on micro-credentials is analyzed according to years, it is found that it followed a relatively stable course between 2015-2020, but it has shown a significant upward trend since 2021. This trend shows that micro-competencies are attracting more and more academic attention, and research in this field is gaining momentum. The sharp rise recorded especially between 2021 and 2024 can be attributed to both the acceleration of digitalization processes on a global scale and the increasing interest of individuals and institutions in lifelong learning and skills-based education models. In this period, the increased adoption of micro-competencies by policy makers, educational institutions and the business world due to their potential to respond more quickly to the dynamic requirements of the labor market has also accelerated the reflection of the issue in the academic field (Çelik & Toprak, 2024).

2024 stands out as the year in which the highest academic research was produced with 29 articles. This shows that academic awareness and research intensity on micro-competencies has reached its peak. However, the main reason for the relatively low number of articles in 2025 is that the data screening process used in this study was completed in February. Therefore, studies to be published during the rest of the year were not included in the analysis.

Distribution of studies by journals

Table 3. Distribution of Studies by Journals

Journal	f
International Journal of Educational Management	4
Education and Information Technologies	3
Distance Education	3
Journal of Continuing Higher Education	3
International Journal of Information and Learning Technology	2
International Journal of Educational Technology in Higher Education	2
Frontiers in Education	2

Innovation in Aging	2
Sustainability	2
Communications of The Association for information Systems	2

Table 3 shows an overview of the 10 most influential journals with articles published on the concept of micro-credentials. International Journal of Educational Management ranks first as the journal with the highest number of articles with 4 articles. Education and Information Technologies, Distance Education, Journal of Continuing Higher Education are in second place with 3 articles each. International Journal of Information and Learning Technology, International Journal of Educational Technology in Higher Education, Frontiers in Education, Innovation in Aging, Sustainability and Communications of the Association for information Systems follow the others with 2 articles each. These journals seem to be specialized in terms of addressing new technologies, management approaches and digital transformation in the field of education. When we consider journals that stand out in the scopes such as online education and distance education, it shows that this concept is closely related to online and distance education. The journals based on higher education level show that micro-credentials studies are mainly conducted within the scope of higher education.

Distribution of Authors by Country

Table 4. Distribution of Authors by Country

Country	f
USA	27
Australia	13
Canada	8
Malaysia	7
Ireland	6
New Zealand	5
England	4
Finland	4
Spain	4
Italy	3
People's Republic of China	3
Poland	3
Cyprus	2
France	2
Japan	2
Pakistan	2
Suudi Arabia	2
Scotland	2
Other	18

An analysis of the distribution of academic studies on micro-credentials by country Table 3 shows that the USA is the largest contributor with 27 articles. The USA is followed by Australia with 13 articles. This suggests that micro-credentials are particularly relevant to educational institutions and policies in the USA and Australia and that the national level of national research recognizes this as a vital area of interest. With 8 articles, Canada comes third; followed by Malaysia (7) and Ireland (6). Research in this field is not confined to particular areas; studies from these nations reflect the geographical variety of academic output on micro-credentials. Micro-credentials are progressively being embraced as a worldwide educational strategy and this issue has become a major focus of research in the international academic community according to contributions from many continents.

Furthermore, the equitable distribution of cross-country cooperation and independent research initiatives shows that the study of micro-credentials is both shaped inside the framework of national policies and handled from an international viewpoint. This offers crucial hints regarding the educational systems and labor markets of micro-credentials positioned in different nations. Furthermore, the cooperation of academics from all around shows that digital learning environments and global education policies are progressively including micro-credentials. Future cross-country partnerships and comparisons of implementation strategies in various environments will help to further enhance the scientific knowledge in the field of micro-credentials.

Distribution by Research Area



Figure 2. Distribution by Research Area

Examining the academic studies on micro competencies in the Web of Science (WOS) database reveals that most of them (n=53) were carried out in the field of educational sciences. According to the results, micro-credentials are now more and more accepted structure in educational systems. The scholarly interest in analyzing the effects of micro-credentials on teaching-learning processes, individual skill development, and lifelong learning policies reflects in the density of studies in educational sciences (Cowie & Sakui, 2021).

Following educational sciences are Computer Science related fields (n=13), Business and Economics (n=6), Engineering (n=5), and Communication (n=3). Micro-credentials are clearly linked to digital learning environments, online education platforms, and technology-supported skill development processes (Berkling, Hänisch & Schütz, 2023), according to the great volume of studies in Computer Science-related domains (n=13). Particularly in the framework of micro-credentials (McGreal, 2023), issues including digital identity verification, blockchain-based certification systems, and AI-supported learning analytics are under increasing attention. With n=6 studies in the field of Business and Economics, the existence of research analyzing the impact of micro-credentials on labor markets, professional development of employees and companies' skills-based recruiting procedures is revealed (Hanafy, 2020). Research in this area has mostly focused on how micro-credentials transform employers' competency assessment strategies and prove individuals' skills independent of conventional diploma systems (Yılık, 2021). Although there are rather few studies in the field of engineering (n=5), it seems that research in this field usually concentrates on issues such the acquisition and certification of technical skills and the function of micro-credentials in engineering education. Research has also focused on engineering students' aspirations toward micro-credentials and the effects of digital badges on their employment (Gregg et al, 2022). The three communication studies show that research on the effects of micro-credentials on digital media, network communities, and interpersonal interaction is still under progress (Keoy et al., 2023). This distribution shows that micro-credentials cover not only education but also in other fields including computer technology, economics, engineering, and communication.

Distribution by Sustainable Development Goals

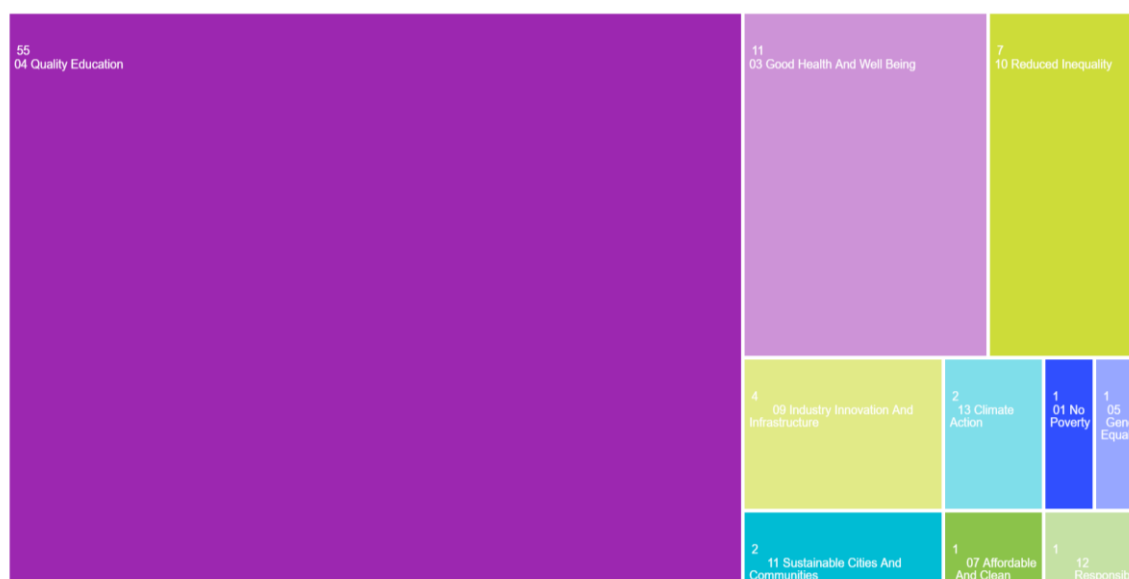


Figure 3. Distribution by Sustainable Development Goals

Analyzing the distribution of studies based on sustainable development goals shows that they are essentially assessed within the purview of quality education (n=55). There are 11 studies addressing "health and wellbeing." Next are "inequality" (n=7), "innovation and infrastructure" (n=4), "sustainability and climate" (n=4).

Analysis of the way academic studies are distributed within the framework of Sustainable Development Goals reveals that most of them are assessed under the name "Quality Education." This can be explained by the fact that one of the pillars of sustainable development, education policies directly influence long-term environmental, social, and economic development as well as directly affect. By raising people's social mobility, education promotes innovation and helps the process of sustainable development, so lowering inequalities.

Conversely, 11 studies fall under the "health and wellbeing" category, which relates to important concerns including public health policies, management of communicable and chronic diseases, and access to health services. Reflecting the scholarly enthusiasm in social justice, equal economic opportunity, and empowerment of underprivileged groups, there are seven studies under the heading "reducing inequalities." In the framework of sustainable development, issues including income disparity, gender inequality, and migrant rights are important. Moreover, the 4 studies falling under the categories "industry, innovation and infrastructure" and "climate action and sustainability" reflect the increasing scholarly interest in mitigating climate change as well as the respect of technological innovation and infrastructure investment as drivers of sustainable development. As a result, micro-credentials, as an important learning approach supporting sustainable higher education, promote academic and professional development by increasing students' employability and enabling them to acquire the necessary skills (Narayanaswamy et al., 2024). Micro-credentials to be provided by educational institutions can be considered as an innovative strategy to support green transformation through sustainability-based digital skills and certification programs.

Distribution by the underlying subjects and concepts

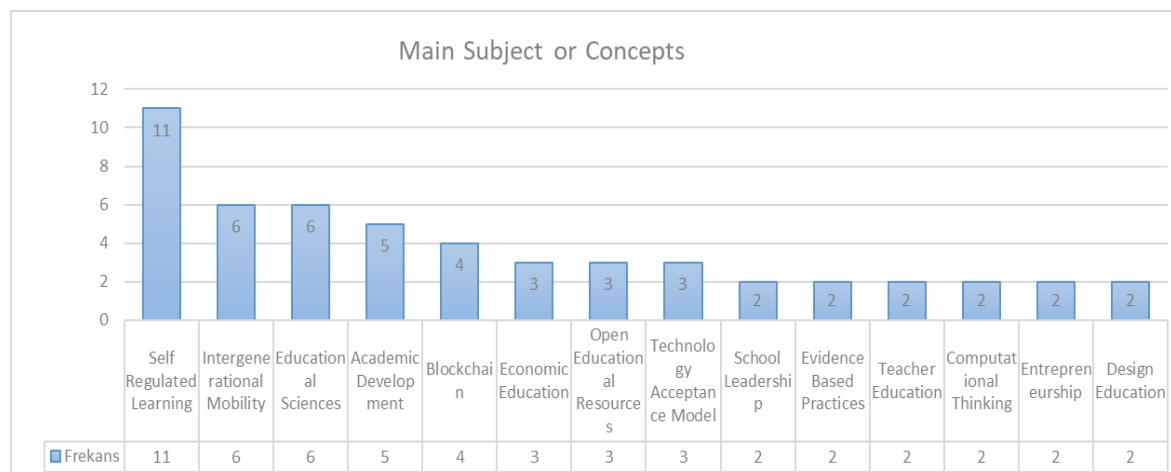


Figure 4. Main subjects and concepts

When the subjects and concepts on which the studies on micro-credentials are based are examined, it is seen that the most emphasized topic is self-regulated learning ($f=11$) (Figure 4). Self-regulated learning is a critical concept for individuals to consciously control their learning processes, set goals and direct their learning processes (Winne, 2016). Various studies have shown that micro-credentials support self-regulated learning processes and increase individuals' responsibility for learning (Du et al., 2015).

Following this, concepts such as intergenerational mobility ($f=6$), academic development ($f=5$), economic education ($f=3$), open educational resources ($f=3$), and technology acceptance model ($f=3$) are seen to be prominent in micro-credentials studies. The concept of intergenerational mobility refers to the capacity of individuals to improve their socioeconomic status through education and the intergenerational transfer of this process. It has been stated that micro-credentials promote professional mobility by providing individuals with flexible and accessible learning opportunities (Bozkurt & Brown, 2022).

In addition, it was found that each of the topics such as school leadership, evidence-based practices, teacher education, computational thinking, entrepreneurship and design education were addressed in two studies. Micro-credentials are found to be a useful tool for professional development and teacher education since they let educators record their classroom activities (Berry, Airhart & Byrd, 2016). Micro-credentials are underlined as extensively used in the field of teacher education in terms of supporting teachers to evaluate their classroom practices with an evidence-based approach (Epstein, Eirich & Wildesen, 2024). Micro-credentials also relate to problems including technology acceptance models and entrepreneurship. Particularly in the framework of entrepreneurship education, studies on the dependability of micro-credentials and the degree of acceptance by stakeholders have exposed that micro-credentials should have well-defined standards to be effective (Phelan & Glackin, 2019).

Micro-credentials studies are thus mostly related in the framework of self-regulated learning, but they also have relevance in many other spheres including academic development, entrepreneurship, technology acceptance model and teacher preparation. This implies that micro-credentials are seen as a tool supporting professional development, academic growth, and socioeconomic mobility as well as a mechanism supporting individual learning processes.

Co-Authorship analysis

According to the co-authorship analysis of the authors, a network map was created by determining at least 1 publication and at least 1 citation criteria to identify the most connected and collaborating authors.

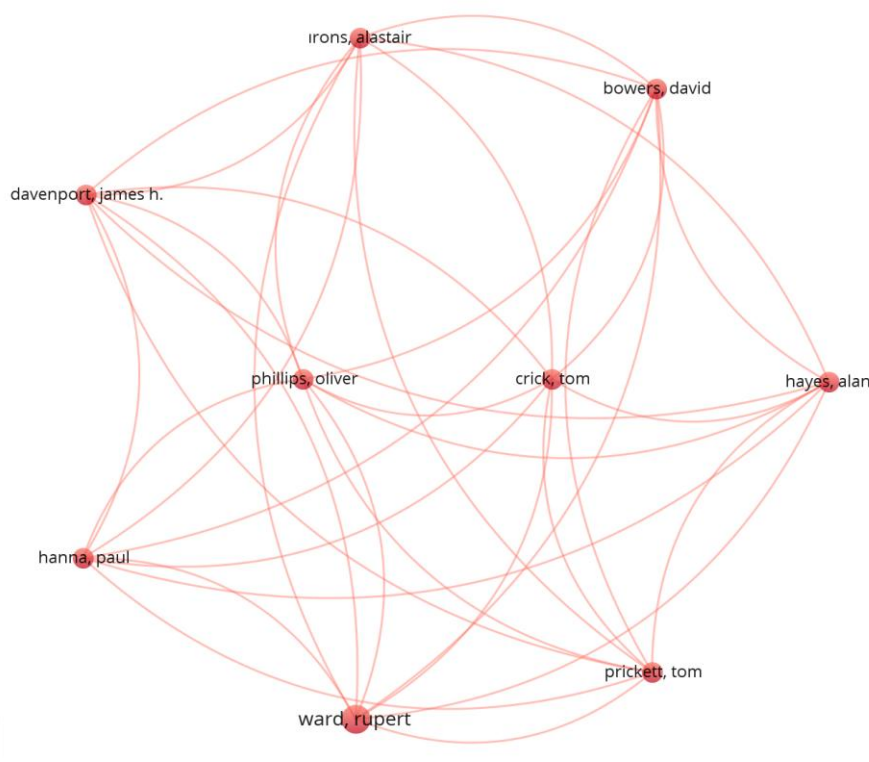


Figure 5. Co-Author links indicating collaboration between authors

Co-author analysis is a technique applied to understand the framework of academic contact between researchers and scientific cooperation. This paper investigates the formation and impact of research groups as well as co-authorship in scientific publications to map cooperative relationships in every academic community (Savič et al., 2019). Often modeled in terms of authors and partnerships, co-authorship networks offer significant new perspectives on the dynamics of academic communities (Jeong et al., 2014).

Under this framework, the study turned up 36 links and 9 authors overall. Every author in the cluster turned out to have 8-unit links. Academic literature claims that highly connected writers have a major influence on knowledge diffusion and often play a central part in academic collaborations (Yang & Zhu, 2015). Sometimes, though, the highly cited study authors might not be among the most connected ones. This implies that independently of academic cooperation, individually significant papers can also be quite highly cited.

With thirty references, the paper by Alsobhi et al. was the most often referenced work in the analysis. These authors were not among the most linked writers, though. According to the literature, some academics can have great influence without having a big network of co-authors (Dehdarirad & Nasini, 2017). Usually, these kinds of situations can be justified by elements like the originality of the topic, the originality of the approach, or the timing of the study.

Citation of authors analysis

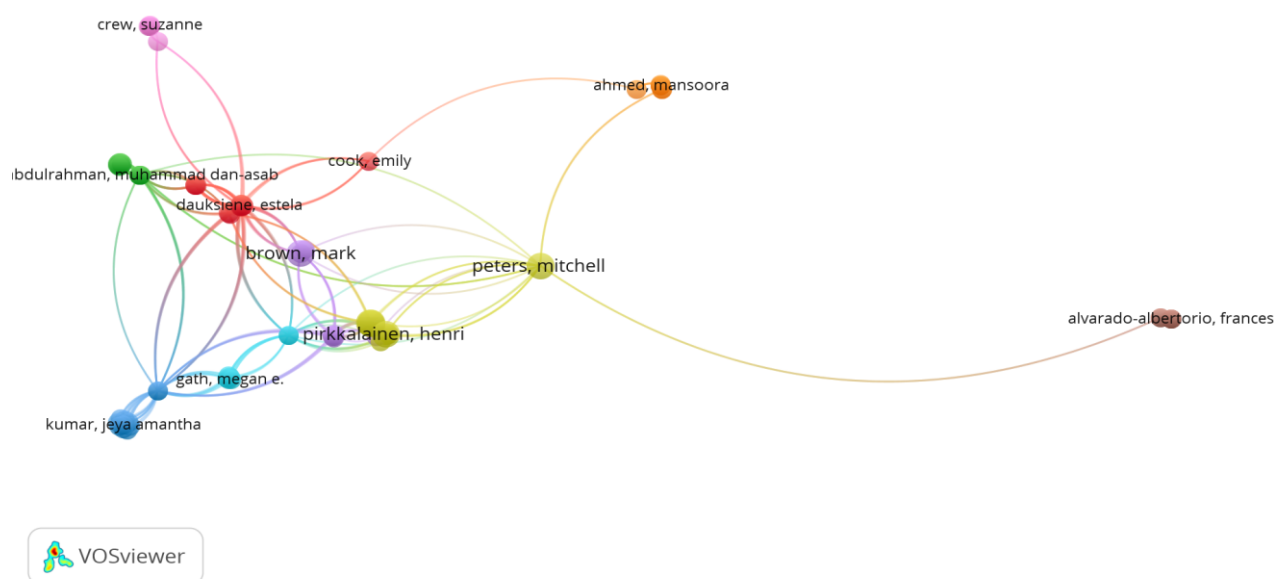


Figure 6. Authors' citation network map

Citation networks help academic literature to map knowledge flow and scientific interactions. This study shows how references in academic publications affect the interactions between researchers and the influence of important studies in scientific domains (McLarr & Bruner, 2022). Often used to ascertain the degree of influence of individual researchers or studies, identify new research trends, and comprehend scientific development in particular fields, citation network analyses are Asatani et al., 2018. In this setting, the analysis produced a citation network map comprising of researchers having at least one publication and at least one citation. 9 clusters, 315 links and a total link strength of 328 were found by means of the analysis conducted on 78 units observed to be related to one another within the scope of the research. Research on comparable citation networks in the literature reveals that they offer crucial information to grasp how academic knowledge is shared and how scientific communities are oriented towards specific focal points (Henrique et al., 2018). With 29 link strengths, Frank Ponte and Christine Rudy were found by the analysis to be the researchers most in influence. Link strength shows the density of citation links (Zhu & Ban, 2018) as well as the central position of a researcher in the academic network. Furthermore, these scholars came out as rather high in terms of citation count. According to similar studies published in the literature, academics with strong links usually have highly cited studies, so enhancing their academic influence (Jung & Segev, 2014).

Understanding the dynamics of scientific cooperation can be much enhanced by citation network analysis. These studies expose the scientific influence of researchers and studies, so clarifying the flow of knowledge (Ganguly & Pudi, 2019). Regarding high link strength and citation counts, researchers like Frank Ponte and Christine Rudy are positioned as significant players in the scientific field. This indicates that knowledge diffusion in the scientific field is directly linked not only to the centrality and connectivity inside the academic network but also to the number of publications.

Citation analysis of countries

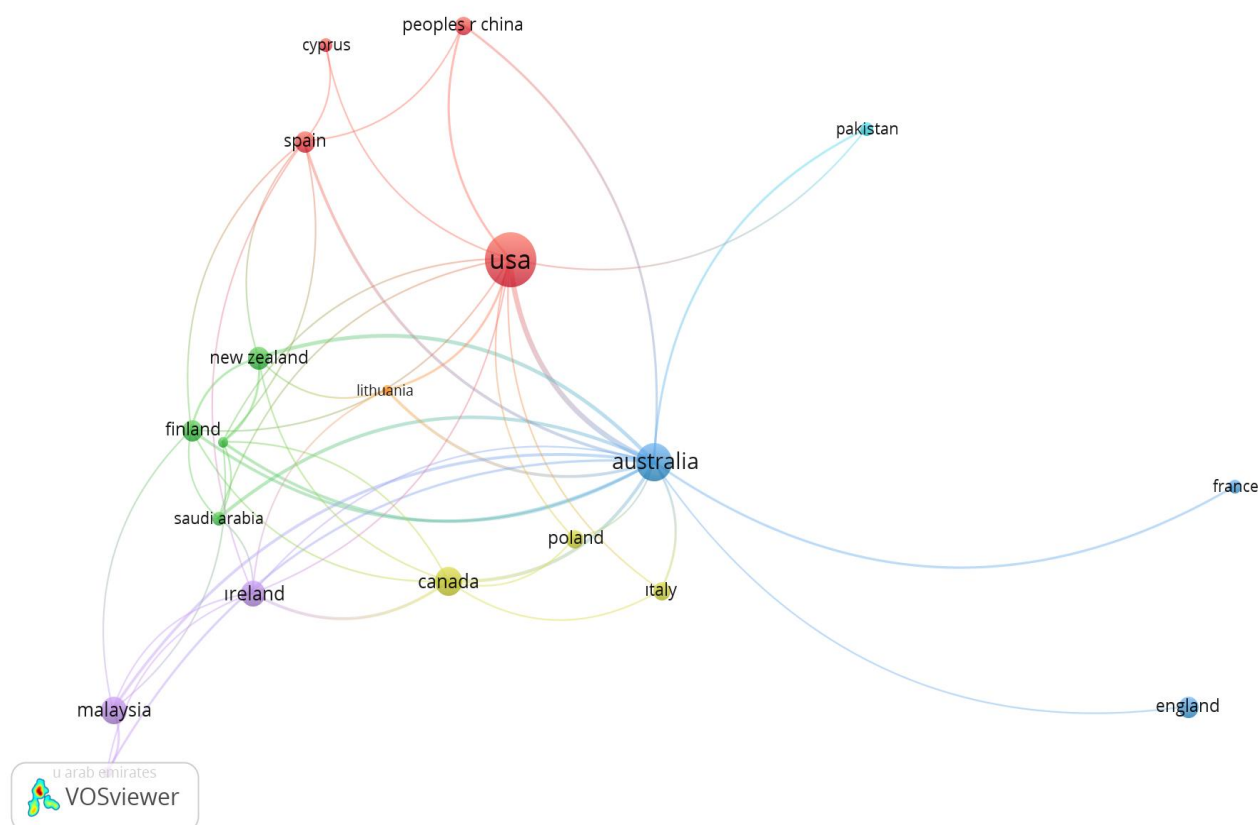


Figure 7. Citation links of countries

Publications, citations, and co-author relationships help to evaluate scientific cooperation across borders (Sadatmoosavi et al., 2018). By mapping the most powerful nations and knowledge transfer routes in each field, this study offers an insight into scientific production and impact. In this study, at least one academic publication and at least one citation were examined using a citation network map constructed over eighteen observation units. This study yielded 7 clusters, 51 links, and 92 overall link strengths (Figure 7). Australia (127 references), USA (94 references), Canada (31 references), and Finland (29 references) are the most often mentioned nations. Australia, USA, Canada, and New Zealand rank highest among total link strength. These results highlight national variations in scientific cooperation and citation performance. Comparable studies published in the literature reveal that countries' academic cooperation, scientific policies and funding strategies directly affect the citation rates of scientific publications (Zhao & Strotmann, 2015). Studies reveal that some nations raise their citation performance by means of more academic cooperation. Regarding scientific cooperation and co-author networks, nations like the US and Australia occupy a central role. Likewise, nations like Finland and Canada seem to have smaller but rather strong networks in terms of academic output. This indicates that the centrality and connectivity inside the scientific community as well as the quantity of publications directly affect citation performance (Chien & Lin, 2018).

Understanding academic interactions and science policies depends much on scientific citation networks spanning nations. Among the nations having the highest citation rates are Australia, the USA, Canada, and Finland; Australia and the USA lead in terms of overall link strength. This demonstrates how closely scientific citation rates correlate with the quality of research as well as with the strength of international academic networks and cooperative projects. Countries trying to boost scientific interactions and foster academic cooperation may find this instructive. Participating in international research networks can be a very important tactic for developing nations trying to raise their citation performance.

Co-Citation analysis of authors

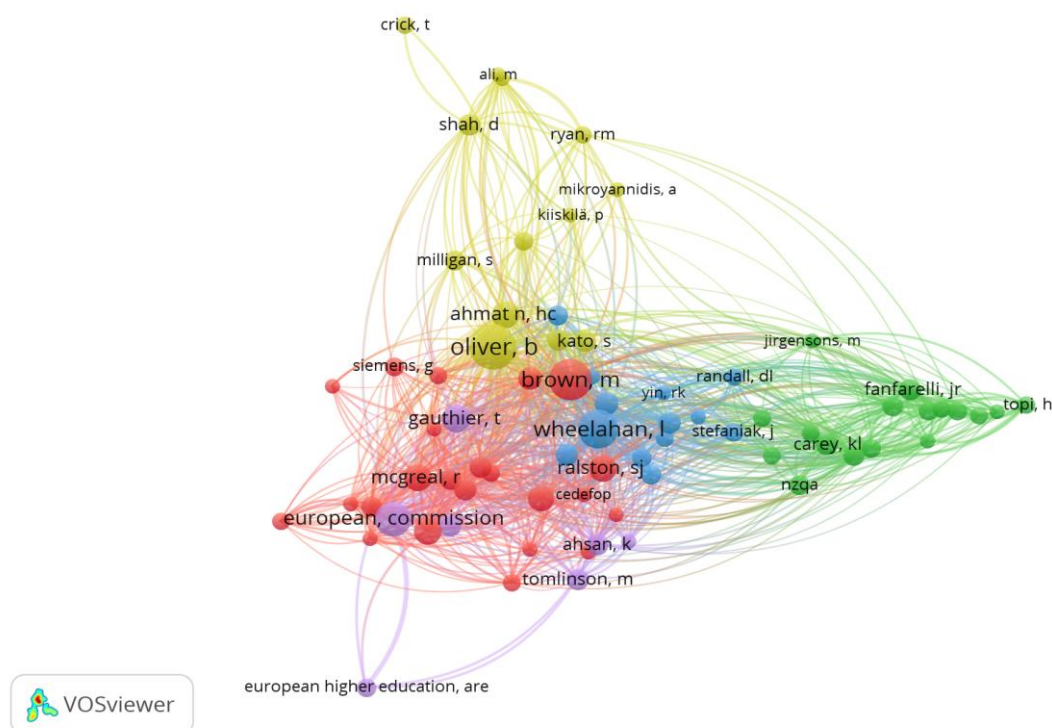


Figure 8. Co-citation analysis of authors

Different sources cited in a publication are called co-citation. According to the analysis conducted on 73 units by selecting a minimum number of 5 citations, 5 clusters and 1470 links are identified in total. The most frequently cited authors are Oliver (41), Brown (34) and Wheelahan (30).

Findings related to keywords

When the most frequently used keywords in the publications related to micro-credentials are analyzed, the keywords micro-credential with 57 repetitions, higher education with 13 repetitions, digital badge with 11 repetitions, employment with 8 repetitions and online learning with 6 repetitions are in the top 5 (Table 5).

Table 5. Keywords

Keywords	f
Micro-credentials	57
Higher Education	13
Digital badge	11
Employment	8
Online learning	6
Skills development	4
Lifelong learning	4
Innovation	3
Teacher Education	3
MOOCs	3
Digital competencies	3
Blockchain	3
Micro-credentials courses	2

An effective approach for spotting research trends, conceptual links, and shared themes in each academic field is keyword analysis (Duvvuru et al., 2013). Higher education, digital badge, employment and lifelong learning were the keywords with the strongest linking power in micro-credential research. 280 observation units were used at least once and the relationships between them were examined in the analysis; the following conclusions were derived: There were found 1021 links and 32 clusters. Higher education, digital badge, employment and lifelong learning have the keywords with strongest link strength. These results reveal that micro-credentials directly relate to increased education and employment possibilities (Salmon, 2023). Though it is not the most often used keyword, the fact that lifelong learning has a high link strength indicates that micro-credentials are vital not only in vocational education but also in the ongoing personal development (Fitzgerald et al., 2022).

The fact that colleges are progressively using digital badges and short-term certificate programs helps one to understand the link of micro-credentials with higher education (Pickard et al., 2018). Conversely, the regular coexistence of digital badges with micro-credentials is connected to the fact that such small-scale certificates are verifiable and shareable (Bozkurt & Brown, 2022). Moreover, the idea of employment and micro-credentials seems to be closely related. Employers are reportedly considering micro-competencies as a crucial tool in demonstrating candidates' technical competencies and are progressively paying such certificates some thought (Booth et al., 2024). The great relevance of the idea of lifelong learning, however, implies that micro-credentials are crucial for both ongoing education and personal development as well as for job entrance (Sargent et al., 2023).

Micro-credentials are most usually connected with the ideas of higher education, digital badge, employment and lifelong learning, according to the keyword network analysis. Though not the most common word, the fact that the idea of lifelong learning has a great connection strength indicates that micro-credentials are extremely important not only in academic and occupational settings but also in the long-term learning processes of people. These results imply that the inclusion of micro-credentials into educational systems should not only be employment-oriented but also regarded as a model that supports the ongoing personal development (Schultz, 2024).

Keyword co-occurrence analysis

Five clusters are found by means of co-occurrence analysis of the publication keywords. These are micro-credentials and employability, learning and skills, distance learning, digitalization and technology, quality assurance and sustainability use.

Cluster 1 is labeled as employability and micro-credentials. Among the ideas in this cluster are those of accreditation, employability, education, human capital theory, lifelong learning, professional development. Micro-credentials are thus seen to be a useful tool for raising people's employability in this setting. From the standpoint of human capital theory, especially, micro-credentials help workers to be more competitive in the labor market by raising their knowledge and skill level. Furthermore, chances for constant professional development and accreditation procedures let people change their career routes and update their competencies in line with the idea of lifetime learning. This helps to support the instruction of adaptable and competent people who can meet the fast-changing needs of the corporate environment.

Cluster 2 is tagged as distance learning. Among these ideas, distance education, MOOC, lifelong learning, flexible learning, educational technology, digital badge, and online learning really stand out within this cluster. In this sense, digital technologies and distance learning models help people's lifelong learning processes by offering flexible learning opportunities. Furthermore, micro-credentials like digital badges and online certificates help people to confirm acquired abilities and get worldwide acceptance. Personalized learning experiences that enable individuals' self-directed learning processes to be more effective pique increasing interest as educational technologies continue to develop.

Cluster 3 is labeled as learning and skills. In the cluster, the concepts of higher education, self-regulated learning, motivation, pre-service teachers, teacher education, skills gap, self-efficacy, learning skills, competencies stand out. In this context, developing individuals' skills to manage their own learning processes (self-regulated learning) and increasing their motivation to learn play a critical role, especially in the context of teacher education and higher education. Micro-credentials help to close the skills gap in education and enable pre-service teachers and instructors to become more competent by supporting their professional development. At the same time, the development of learning skills and competencies stands out as one of the key elements that enable individuals to succeed in both their academic and professional careers.

Cluster 4 is labeled as digitalization and technology. In this cluster, the concepts of technology, blockchain, digital credentials, digital badges, open badge, computer science education, and innovation are prominent. In this context, the digitalization process is radically changing education systems, and the validation of qualifications, especially blockchain-based solutions and digital identification methods enable the recognition of micro-credentials in a secure and visible way. Open badges and digital certificates allow individuals to prove and share their skills on an international scale, while computer science education and innovation ecosystems are among the key elements supporting this transformation. These technology-driven developments accelerate digital transformation in education by making learning processes more flexible, accessible and personalized.

Cluster 5 is labeled as quality assurance and sustainability. In the cluster, the concepts of sustainability, public health, quality assurance, cultural safety, and quality education are prominent. In this context, quality assurance in education contributes to the long-term well-being of individuals and societies by aligning with sustainable development goals. Cultural safety supports inclusion and diversity in education, ensuring that individuals have equal opportunities in different socio-cultural contexts. In addition,

public health and quality education enable individuals to develop in a healthy and sustainable environment in lifelong learning processes and contribute to their professional and personal development through micro-credentials. Ensuring quality assurance in education stands out as a critical factor supporting not only individual success but also social sustainability and economic development.

CONCLUSION AND RECOMMENDATIONS

This study presents a bibliometric analysis of academic studies on micro-credentials, revealing research trends, key themes, and scientific collaborations in the field. The number of academic publications on micro-credentials has increased significantly, especially since 2021. This trend is directly related to the acceleration of digitalization, increasing lifelong learning requirements and the need to adapt to the dynamic requirements of the labor market. It is seen that studies on micro-credentials are mostly published in journals focused on educational sciences. However, there are also studies on this subject in journals related to distance education, digital learning technologies and sustainable education policies. Higher academic contributions on micro-credentials have been found to originate from nations including the USA, Australia, and Canada. These nations have included micro-credentials into their national education strategies and into their policies of instruction. Furthermore, observed is the significance of international cooperation, although in some nations the studies are more locally oriented. Though most of the studies in the field of educational sciences have been carried out, research on micro-credentials has also been done in computer sciences, business, economics, and engineering. This indicates that micro-credentials are a significant problem not only in the framework of education but also in terms of labor market and technology evolution.

Sustainable development goals indicate that most of the studies under the heading Quality Education are related to issues like health and well-being, so lowering inequality and sustainability; yet their titles reflect this concentration. This result reveals that micro-credentials are regarded as an educational tool supporting social and personal growth. Self-regulated learning, intergenerational mobility, academic development and digital badges have emerged as the most often mentioned ideas within the parameters of micro-credential studies. Furthermore, it is seen that topics such as entrepreneurship, technology acceptance model and teacher training also draw attention to academic interest.

Keyword analysis revealed that micro-credentials have strong connections to higher education, digital badges, employment and lifelong learning. At the same time, it is determined that the most cited countries and researchers are mostly from countries such as the USA, Australia and Canada. As a result, micro-credentials is gaining more and more importance as a flexible and innovative learning model that responds to the needs of modern education systems and labor markets. However, addressing this area more comprehensively, encouraging interdisciplinary collaborations and applied research is a critical requirement to maximize the impact of micro-credentials in education and business.

Recommendations

- Addressing micro-credentials not only from the perspective of educational sciences but also from different disciplines such as economics, engineering, computer sciences and social sciences will enable the subject to be evaluated from a broader perspective.
- While countries such as the USA, Australia and Canada have high academic productivity in the field of micro-credentials, increasing joint studies with developing countries will contribute to making the knowledge in this field more inclusive.
- More empirical research is required to examine the effects of micro-credentials on professional development, employment and career development. Such research can guide education policies and labor force planning.
- Standards can be developed to increase the recognition and reliability of verifiable micro-credentials such as digital badges and block chain-based certificates. This can help ensure wider acceptance among employers and educational institutions.
- The role of micro-credentials in education systems should be addressed further in line with sustainable development goals. How micro-credentials can be used more effectively in terms of reducing inequalities in education and supporting lifelong learning processes should be investigated.
- Integrating micro-credentials into teacher training programs can support teachers' professional development and increase their capacity to provide skill-focused education to students. In this context, more experimental studies should be conducted in the field of teacher training.

Researchers' contribution rate

The research process's theoretical framework, methodology, data collection, data analysis, evaluation of the findings, and manuscript drafting were carried out by the first and only author.

Statements of publication ethics

I hereby declare that the study has not unethical issues and that research and publication ethics have been observed carefully.

Ethics Committee Approval Information

This study does not require approval from the ethics committee.

Funding

The author received no financial support for the research, author-ship, and/or publication of this article.

REFERENCES

- Ahmat, N. H. C., Bashir, M. A. A., Razali, A. R., & Kasolang, S. (2021). Micro-credentials in higher education institutions: Challenges and opportunities. *Asian Journal of University Education*, 17(3), 281-290.
- Ali, S. A., & Khan, R. (2023). Harnessing the digital nature of Micro-Credentials in Higher Education. *International Journal of Education and Teaching*, 3(1), 58-66.
- Asatani, K., Mori, J., Ochi, M., & Sakata, I. (2018). Detecting trends in academic research from a citation network using network representation learning. *PLoS ONE*, 13(5), e0197260. <https://doi.org/10.1371/journal.pone.0197260>
- Berry, B., Airhart, K. M., & Byrd, P. A. (2016). Microcredentials: Teacher learning transformed. *Phi Delta Kappan*, 98(3), 34-40.
- Bozkurt, A., & Brown, M. (2022). Microcredentials: Stackable, combinable, or transferable qualifications. In *EdTechnica: The Open Encyclopedia of Educational Technology*. EdTech Books. <https://edtechbooks.org/encyclopedia/microcredentials>
- Çelik, G. & Torak, E. (2024). Mikro Yeterliliklerin Kredilendirilmesinde Kalite Güvencesi ve Akreditasyon Konusunda Öne Çıkan Yaklaşımlar ve Gelecek Senaryoları. *4th International Conference on Educational Technology and Online Learning – ICETOL 2024*, 566-573, Balıkesir, Türkiye.
- Chien, B. C., & Lin, K. Y. (2018, July). Analysis of retrieving publications with multiple citation networks. In *Proceedings of the 5th Multidisciplinary International Social Networks Conference* (pp. 1-6).
- Cowie, N., & Sakui, K. (2021). Micro-credentials. *Pacific Journal of Technology Enhanced Learning*, 3(1), 27-28.
- Dehdarirad, T., & Nasini, S. (2017). Research impact in co-authorship networks: A two-mode analysis. *Journal of Informetrics*, 11(2), 371-388.
- Donthu, N., Kumar, S., Mukherjee, D., Pandey, N., & Lim, W. M. (2021). How to conduct a bibliometric analysis: An overview and guidelines. *Journal of Business Research*, 133, 285-296.
- Du, Z., Wang, X., Ma, H., & Liang, D. (2015, November). Research on Self-regulated Learning Platform based on Massive Micro-Lessons. In *2015 International Conference on Social Science, Education Management and Sports Education* (pp. 1275-1278). Atlantis Press.
- Duvvuru, A., Radhakrishnan, S., More, D., Kamarthi, S., & Sultornsanee, S. (2013). Analyzing structural & temporal characteristics of keyword system in academic research articles. *Procedia Computer Science*, 20, 439-445.
- Epstein, J., Eirich, J., & Wildesen, J. (2024). Engaging teacher's mental health self-care early through a trauma-informed microcredential. *School-University Partnerships*, 17(2), 153-169.
- Fitzgerald, R., Huijser, H., Umarova, A., & Adachi, C. (2022). Reconnecting lifelong learners through short courses and microcredentials designed with industry partners. *ASCILITE Publications*, e22141-e22141.
- Ganguly, S. (2019). *Exploiting Textual Content in Academic Citation Networks (Doctoral dissertation)*. International Institute of Information Technology, Hyderabad.
- Gregg, A., Park, J., Fenton, C., Lang, D., & Handley, M. (2022, October). Exploring the “Why” of Micro-Credentials and Digital Badges: Engineering Students’ Motivations for and Perceived Utility of Learning Outside of Class. In *2022 IEEE Frontiers in Education Conference (FIE)* (pp. 1-7). IEEE.
- Hanafy, A. (2020). *Features and Affordances of Micro-Credential: Platforms in Higher Education (Master Thesis)*. Tampere University, Finland.
- Henrique, B. M., Sobreiro, V. A., & Kimura, H. (2018). Building direct citation networks. *Scientometrics*, 115, 817-832.
- Jeong, Y. K., Song, M., & Ding, Y. (2014). Content-based author co-citation analysis. *Journal of Informetrics*, 8(1), 197-211.
- Jung, S., & Segev, A. (2013, October). Analyzing future communities in growing citation networks. In *Proceedings of the 2013 international workshop on mining unstructured big data using natural language processing* (pp. 15-22).
- Kayyali, M. (2025). Defining Micro-Credentials: Concepts, Theories, and Pedagogies. In *Navigating Quality Assurance and Accreditation in Global Higher Education* (pp. 219-246). IGI Global Scientific Publishing.
- Keoy, K. H., Iqbal, J., Ho, M. C., Ooi, P. B., Anjum, S. S., Koh, Y. J., ... & Lillian, S. Y. K. (2023, July). Preliminary Exploratory Study: Are HEPs Ready for Communication & Multimedia Micro Credentials Adoption in Malaysia?. In *2023 International Conference on Digital Applications, Transformation & Economy (ICDATE)* (pp. 1-5). IEEE.
- Kır, Ş., & Bozkurt, A. (2022). Yükseköğretimde Mikro-Krediler Üzerine Kavramsal Bir Değerlendirme. *Uşak Üniversitesi Eğitim Araştırmaları Dergisi*, 8(1), 12-29.

- Kukkonen, A. (2021). *The value of microcredentials for higher education students* (Unpublished Doctoral dissertation). Tampere University, Tampere.
- McGreal, R. (2023). Blockchain and Micro-Credentials in Education. *International Journal of E-Learning & Distance Education*, 38(1), 1-17.
- McGreal, R., & Olcott, D. (2022). A strategic reset: Micro-credentials for higher education leaders. *Smart Learning Environments*, 9(1), 9.
- McLaren, C. D., & Bruner, M. W. (2022). Citation network analysis. *International Review of Sport and Exercise Psychology*, 15(1), 179-198.
- Murgatroyd, S. (2022). "Reinventing" Higher Education—Rhetoric versus Reality in a Post-Pandemic World. *Revista Paraguaya de Educación a Distancia (REPED)*, 3(2), 48-57.
- Narayanaswamy, R., Albers, C. S., Knotts, T. L., & Albers, N. D. (2024). Sustaining and Reinforcing the Perceived Value of Higher Education: E-Learning with Micro-Credentials. *Sustainability*, 16(20), 8860.
- Orman, R., Şimşek, E., & Çakır, M. A. K. (2023). Micro-credentials and reflections on higher education. *Higher Education Evaluation and Development*, 17(2), 96-112.
- Page, M. J., Moher, D., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., ... & McKenzie, J. E. (2021). PRISMA 2020 explanation and elaboration: updated guidance and exemplars for reporting systematic reviews. *BMJ*, 372:n71.
- Phelan, S. E., & Glackin, C. E. (2021). A cautionary note on microcredentialing in entrepreneurship education. *Entrepreneurship Education and Pedagogy*, 4(3), 261-281.
- Pickard, L., Shah, D., & De Simone, J. J. (2018, September). Mapping microcredentials across MOOC platforms. In *2018 learning with MOOCs (LWMOOCs)* (pp. 17-21). IEEE.
- Pirkkalainen, H., Sood, I., Padron Napoles, C., Kukkonen, A., & Camilleri, A. (2023). How might micro-credentials influence institutions and empower learners in higher education?. *Educational Research*, 65(1), 40-63.
- Raj, R., Singh, A., Kumar, V., & Verma, P. (2024). Achieving professional qualifications using micro-credentials: a case of small packages and big challenges in higher education. *International Journal of Educational Management*, 38(4), 916-947.
- Rižnar, I. (2023). Navigating micro-credentials: An analysis of student needs, perceptions and skills development at a Slovenian university. *E-mentor*, 101(4), 63-70.
- Sadatmoosavi, A., Nooshinfard, F., Hariri, N., & Esmaeil, S. M. (2018). Does the superior position of countries in co-authorship networks lead to their high citation performance. *Malaysian Journal of Library and Information Science*, 23(1), 51-65.
- Salmon, M. (2023). Drivers of the global push for microcredentials in higher education: Flexibility and employability in contemporary university systems. *Perspectives: Policy and Practice in Higher Education*, 27(4), 179-187.
- Sargent, J., Rienties, B., Perryman, L. A., & FitzGerald, E. (2023). Investigating the Views and Use of Stackable Microcredentials within a Postgraduate Certificate in Academic Practice. *Journal of Interactive Media in Education*, 2023(1), 1-12.
- Savić, M., Ivanović, M., Jain, L. C., Savić, M., Ivanović, M., & Jain, L. C. (2019). Co-authorship networks: an introduction. *Complex networks in software, knowledge, and social systems*, 179-192.
- Schultz, C. (2024). Experiences of microcredentials in Canada: perspectives, possibilities, and a case for national data collection. *International Review of Applied Economics*, 38(4), 410-421.
- Tan, K. T., Chan, J. I. L., Arumugam, P. V., & Lee, H. W. (2024). Closing The Talent Gap: A Proposed Micro-Credential Model In Malaysian Formal Education. *Turkish Online Journal of Distance Education*, 25(4), 67-80.
- Teksöz, G. (2014). Geçmişten ders almak: Sürdürülebilir kalkınma için eğitim. *Bogazici University Journal of Education*, 31(2), 73-97.
- Van Eck, N. J., & Waltman, L. (2014). Visualizing bibliometric networks. In *Measuring scholarly impact: Methods and practice* (pp. 285-320). Cham: Springer International Publishing.
- Ward, R., Crick, T., Hanna, P., Hayes, A., Irons, A., Miller, K., ... & Walters, J. (2023). Using skills profiling to enable badges and micro-credentials to be incorporated into higher education courses. *Journal of Interactive Media in Education*, 2023(1), 1-22.
- Winne, P. (2016). Self-regulated learning. *SFU Educational Review*, 9, 535-540. <https://doi.org/10.1016/B978-0-08-097086-8.25091-5>
- Yang, Z., & Zhu, C. (2015, October). Determine the Influence of Academic Research Base on Social network analysis. In *2015 3rd International Conference on Mechatronics and Industrial Informatics (ICMII 2015)* (pp. 740-744). Atlantis Press.
- Zhao, D., & Strotmann, A. (2015). *Analysis and visualization of citation networks*. Morgan & Claypool Publishers.
- Zhu, X. P., & Ban, Z. (2018, May). Citation count prediction based on academic network features. In *32nd international conference on advanced information networking and applications (AINA)* (pp. 534-541). IEEE.



| Research Article / Araştırma Makalesi |

The Effect of Argumentation-Based Practices on 8th Grade Students' Views on Socioscientific Issues

Argümantasyon Temelli Uygulamaların 8. Sınıf Öğrencilerinin Sosyobilimsel Konulara İlişkin Görüşlerine Etkisi

Dilek KARIŞAN¹, Serkan AYDIN²

Keywords

1. Argumentation
2. Socioscientific Issues
3. Science Education
4. Scientific Literacy

Anahtar Kelimeler

1. Argümantasyon
2. Sosyobilimsel Konular
3. Fen Eğitimi
4. Fen Okuryazarlığı

Received/Başvuru Tarihi
02.03.2025

Accepted / Kabul Tarihi
10.10.2025

Abstract

Purpose: This study aims to examine the impact of argumentation-based practices on eighth-grade students' perspectives on socioscientific issues.

Design/Methodology/Approach: The research employed a case study design, one of the qualitative research methods. The study group consists of 31 students enrolled in a public school affiliated with the Ministry of National Education in Aydın. Participants were selected using a purposive sampling method. Data were collected through semi-structured interviews, classroom discussions, video recordings, and students' written worksheets. For data analysis, the argumentation assessment rubric developed by Sadler and Fowler (2006) was utilized.

Findings: The findings indicate that integrating socioscientific issues with argumentation positively influences students' perspectives on these topics.

Highlights: Argumentation-based approaches enhance students' understanding of socioscientific issues. Socioscientific argumentation fosters critical thinking and decision-making skills. Classroom implementation of argumentation can improve students' engagement with complex scientific and social topics.

Öz

Amaç: Bu çalışma, argümantasyon temelli uygulamaların sekizinci sınıf öğrencilerinin sosyobilimsel konulara yönelik bakış açıları üzerindeki etkisini incelemeyi amaçlamaktadır.

Tasarım/Yöntem/Uygulama: Araştırmada, nitel araştırma yöntemlerinden biri olan durum çalışması deseni kullanılmıştır. Çalışma grubunu, Aydın ilinde Milli Eğitim Bakanlığı'na bağlı bir devlet okulunda öğrenim gören 31 öğrenci oluşturmaktadır. Katılımcılar, amaçlı örnekleme yöntemiyle belirlenmiştir. Veriler, yarı yapılandırılmış mülakatlar, sınıf içi tartışmalar, video kayıtları ve öğrencilerin yazılı olarak sundukları çalışma kâğıtları aracılığıyla toplanmıştır. Verilerin analizinde, Sadler ve Fowler (2006) tarafından geliştirilen argümantasyon değerlendirme rubriği kullanılmıştır.

Bulgular: Araştırma bulguları, sosyobilimsel konuların argümantasyon yöntemiyle işlenmesinin öğrencilerin bu konulara yönelik bakış açılarını olumlu yönde etkilediğini göstermektedir.

Önemli vurgular: Argümantasyon temelli yaklaşımlar, öğrencilerin sosyobilimsel konulara yönelik anlayışını geliştirir. Sosyobilimsel argümantasyon, eleştirel düşünme ve karar verme becerilerini destekler. Sınıf ortamında argümantasyon uygulamaları, öğrencilerin bilimsel ve toplumsal konulara ilgisini artırabilir.

1 Corresponded Author, Aydın Adnan Menderes University, Faculty of Education, Department of Science Education, Aydın, Türkiye; <https://orcid.org/0000-0002-1791-9633>

2 Milli Eğitim Bakanlığı, Aydın Efeler İlçe Milli Eğitim, Aydın, Türkiye; <https://orcid.org/0000-0001-7357-2729>

INTRODUCTION

Science and technology, which have been constantly developing from past to present, have also caused the social structure to change. The change from the hunter society structure in the first civilizations to today's society structure with science and technology is a result of this change (Çalık & Çınar, 2009). With the invention of fire, human beings gained features such as warming, enlightenment, protection and cooking, and with the invention of the wheel, they found the opportunity to carry their loads more easily and started to travel farther. The further acceleration of technological developments with the realization of the industrial revolution accelerated the change in the social structure. Since the 21st century includes the fourth industrial revolution, science and technology are developing faster than ever before (Babacan, 2017). This acceleration of developments in science and technology also rapidly changes the needs of individuals. Individuals need to keep up with changes in order to survive and adapt to the developing world. This adaptation process is supported by the knowledge and skills individuals acquire through education (Karaca, 2018; Karişan, 2011). In a qualified education process, individuals are expected to acquire skills such as thinking, criticizing, questioning, researching, making informed choices, and transferring the knowledge acquired to daily life (Alkın-Şahin et al., 2014; Meral, 2018). In order to have these skills, individuals need to analyze, research, critically evaluate and question the events they encounter (Meral, 2018). Today, with the rapid development of technology, access to information has become quite easy. However, although individuals can access so much information, they may accept the information they obtain without questioning and may have difficulty in choosing the right information.

In today's rapidly evolving world of science and technology, fostering science literacy among all individuals has been established as a fundamental goal. Science literacy empowers individuals to comprehend scientific concepts and advancements, equipping them with the knowledge and skills necessary to navigate an increasingly complex world (Başar, 2018). Science-literate individuals are those who actively research, question, and solve problems, make informed and effective decisions, communicate proficiently, collaborate effectively, embrace lifelong learning, think analytically, and approach issues with a reflective and critical mindset (MoNE, 2013). As a multidimensional competency, science literacy encompasses cognitive, affective, and psychomotor skills related to science (Durant, 1993). Furthermore, science literacy enables individuals to understand, track, internalize, and apply scientific and technological advancements, allowing them to utilize scientific knowledge and innovations effectively in their daily lives. By fostering a habit of scientific thinking, it equips individuals with the skills and competencies necessary to make informed choices and enhance their quality of life (Özdemir, 2010). Science, society, and the environment are interconnected and continuously influence one another. Societal needs have driven the emergence and advancement of science. However, scientific developments are not always received uniformly by all members of society. This disparity in perspectives often leads to disagreements and societal divisions. To describe the societal dilemmas that influence science, the concept of "socioscientific issues" has been introduced (Sadler, 2004; Sadler & Zeidler, 2004). Integrating socioscientific issues into science education allows students to recognize real-life problems, critically analyze them, generate possible solutions, and bridge the gap between school learning and everyday experiences (Sevgi, 2016).

Socioscientific issues are topics that encompass both scientific and societal dimensions, lack a single definitive answer, are open to debate, and can be evaluated from multiple perspectives, often involving ethical and moral considerations (Sadler & Zeidler, 2005; Sadler & Fowler, 2006; Zeidler & Nichols, 2009). Examples of such issues include cloning, genetically modified organisms (GMOs), chemical pesticides, global warming, flu vaccines, stem cell research, alternative fuels, and nuclear energy (Birdal, 2019; Karamanlı, 2019; Levinson, 2006). For a topic to be classified as a socioscientific issue, it must be grounded in scientific principles and hold significance in social life (Eastwood et al., 2012). Sadler and Zeidler (2005) emphasize that socioscientific issues also have ethical and moral dimensions, making them crucial for fostering societal character development and virtues.

Given the complex and multifaceted nature of socioscientific issues, individuals must engage in thoughtful reasoning and informed decision-making when addressing them. One of the key methods for navigating these issues is argumentation. Argumentation is defined as a process in which individuals present their goals and interpretations in a logical and structured manner (Patronis et al., 1999). The Turkish Language Association defines the term "argument" as "claim, evidence, assertion, thesis." Therefore, in argumentation, claims and explanations must be supported by evidence (Türköz, 2019).

Through argumentation, individuals propose solutions to real-life problems. To effectively present their arguments and refute counterarguments, individuals must have a strong command of the topic, think critically about the problem, consider multiple perspectives, possess strong communication skills, and provide compelling evidence to persuade others (Prain, 2006). Examining the characteristics of individuals with argumentation skills reveals that this method fosters higher-order thinking and promotes long-term, meaningful learning.

Importance of the Research

The importance of this research lies in its alignment with the rapid advancements in science and technology and their transformative effects on social structures. As science and technology have progressed from early human inventions like fire and the wheel to the acceleration brought by the fourth industrial revolution, they have continuously reshaped society and the needs of individuals (Babacan, 2017; Çalık & Çınar, 2009). This fast-paced development creates a need for individuals to adapt, which is facilitated through education that equips them with skills such as critical thinking, problem-solving, and making informed decisions (Alkın-Şahin et al., 2014; Karaca, 2018). In an era of easy access to vast amounts of information, the ability to critically evaluate and select reliable knowledge is more important than ever (Meral, 2018).

This research is significant as it addresses the critical need to foster science literacy in students by focusing on socioscientific issues topics that blend scientific knowledge with societal concerns, such as global warming, genetically modified organisms (GMOs), nuclear energy, and others. Through the use of argumentation, a method that enhances critical thinking, students are better able to analyze problems, engage in informed discussions, and make reasoned decisions. The study aims to develop higher-order thinking skills in students and equip them with the ability to engage with real-world issues, promoting informed, responsible citizenship. Additionally, this research highlights the importance of integrating socioscientific issues with argumentation in science education, especially at the 8th-grade level, where there is a noticeable gap in the literature regarding this integration. The findings of this study are expected to contribute to the growing body of research in this field, helping to enhance science literacy and the development of scientifically informed citizens capable of making thoughtful decisions in a rapidly evolving world.

Purpose of the Research

The purpose of this research is to investigate the impact of argumentation-based practices on 8th-grade students' perspectives on socioscientific issues. Specifically, this study aims to explore how engaging students in discussions on various socioscientific topics such as endangered species, nuclear energy, global warming, and others can influence their ability to critically analyze these issues, develop informed opinions, and effectively communicate their viewpoints. By integrating argumentation as a method for decision-making, the research seeks to enhance students' skills in reasoning, problem-solving, and making evidence-based decisions. Additionally, the study aims to examine how argumentation helps students bridge the gap between scientific knowledge and real-world applications, fostering science literacy and promoting thoughtful, ethical decision-making in the context of societal challenges. Furthermore, the study seeks to explore the impact of argumentation-based practices on students' perspectives, particularly in how they engage with and understand current socioscientific issues through scientific knowledge and real-world applications.

Research question: What are the opinions of 8th-grade students on discussing current socioscientific issues through the argumentation method following argumentation-based practices?

METHOD

In this study, qualitative methods were used to examine the impact of argumentation-based practices on students' perspectives regarding socioscientific issues. Qualitative research is a methodological approach that focuses on understanding and interpreting the meaning of individuals' experiences, perspectives, and behaviors. This type of research emphasizes exploring complex phenomena within their natural settings, providing rich and detailed insights into specific contexts (Creswell, 2013).

The case study research method guided this study, which is an in-depth, contextually rich investigation of a particular phenomenon within a real-life setting (Yin, 2018). Case study research is particularly valuable when exploring new, complex, or under-researched areas, providing a detailed understanding of the subject in question. In this study, the case study approach allowed the researcher to analyze how argumentation practices influenced the students' argumentation skills and their views on socioscientific issues.

Participants

The participants in this study were 31 eighth-grade students from a public school in Aydın, Turkey. Their backgrounds were diverse, encompassing a wide range of academic abilities, interests, and learning styles. The group included students who were highly engaged in science and showed a natural curiosity about societal issues, as well as those who appeared less motivated and often needed additional encouragement to participate in class discussions. Many of the students came from families with varying levels of education, which influenced their prior knowledge and understanding of science-related topics. While some students had an advanced understanding of environmental issues and were able to discuss topics like climate change, GMOs, and renewable

energy with a high degree of insight, others struggled to grasp the complexities of these issues and often relied on simple, factual explanations.

To deepen the insights into how argumentation skills were developed, purposive sampling was used to select groups for face-to-face interviews. Purposive sampling is a targeted approach often used in qualitative research to focus on specific cases that represent extreme or particularly informative instances related to the research problem. This method enabled the researcher to gain a more nuanced understanding of the variables at play by selecting participants who exhibited rich and varied levels of argumentation (Yıldırım & Şimşek, 2018). Specifically, the study aimed to investigate the perspectives of three distinct groups based on the quality of their arguments: high-level, moderate-level, and low-level argument producers. By focusing on these extreme cases, the study sought to capture the breadth of students' argumentative development and provide a detailed examination of how different levels of argumentation manifested in the context of socioscientific issues.

In terms of their interaction with the content, students exhibited a range of engagement during argumentation-based activities. Some actively participated, presenting well-structured arguments, providing evidence, and engaging in respectful debates with their peers. Others, however, were more reserved, often waiting for their classmates to take the lead in discussions. Despite these differences, over the course of the study, many students showed significant improvement in their ability to formulate arguments, ask critical questions, and consider multiple perspectives. These changes were particularly evident in their increased confidence and willingness to engage in discussions that required deeper reasoning and reflection.

The students' personal engagement with the socioscientific issues discussed in class also varied. For some, topics like the use of pesticides or nuclear energy sparked strong emotional reactions, leading them to actively seek out information and challenge existing viewpoints. For others, these issues seemed distant, and their engagement was more passive, focusing on memorizing facts rather than critically engaging with the material. However, the structured nature of the argumentation activities allowed students to gradually build their argumentative skills, transforming their initial passive engagement into more active participation. The influence of the classroom environment, teacher guidance, and peer interactions also played a crucial role in shaping how the students approached the socioscientific topics and how effectively they developed their argumentation skills.

This thick description highlights not only the academic profiles of the students but also their personal motivations, the socio-cultural context that influenced their learning, and how these factors interacted to shape their experiences during the study. Through this detailed portrayal, the study provides a nuanced understanding of how students' backgrounds, their engagement with the content, and the learning environment contributed to the development of their argumentation abilities.

Study Context

This study was conducted over a 14-week period and aimed to investigate students' argumentation skills through the integration of Toulmin's model of argumentation in socioscientific contexts. The procedure consisted of several phases, including instruction, argumentation activities, data collection, and evaluation. In the initial phase, students were introduced to Toulmin's model of argumentation and engaged in sample activities to familiarize themselves with the structure of claims, evidence, and rebuttals. This preparatory phase lasted four weeks (eight class hours), ensuring that students had a solid understanding of the model before engaging in structured argumentation tasks. Following this, the main argumentation phase began, where students participated in discussions on predetermined socioscientific issues. Each session started with a short video, cartoon, or newspaper presentation providing engaging and relevant background information. This was followed by structured discussions in which students expressed their ideas, challenged opposing viewpoints, and attempted to persuade their peers by refuting counterarguments. Ample space and time were given for students to engage in argumentation-based discourse.

Eight student-generated activity sheets covering topics such as Endangered Species, Weight Loss Drugs, Geothermal Energy, Nuclear Energy, Base Stations, Agricultural Pesticides, Genetically Modified Organisms, and Climate Change were utilized. Students were encouraged to apply Toulmin's model throughout the discussions, and they documented key points. Each activity was discussed over a one-week period, with discussions being video- and audio-recorded for further analysis. Data collection involved multiple methods to assess students' argumentation skills and development over time. Pre- and post-interviews were conducted to examine students' initial perspectives and any changes in their reasoning abilities after the intervention. In addition, written argumentation responses were evaluated using the argumentation assessment rubric developed by Sadler and Fowler (2006).

To gain deeper insights, post-interviews were conducted with three students representing different proficiency levels high, middle, and low argumentation performance on a voluntary basis. These interviews were audio-recorded and analyzed to understand how students' argumentation skills evolved throughout the study. The recorded discussions, written argumentation assessments, and interview data provided a comprehensive understanding of students' engagement with socioscientific issues and the effectiveness of Toulmin's model in fostering structured argumentation skills.

Data collection tools

The data for this study were collected using three main tools: activity worksheets, classroom discussion videos, and face-to-face interview recordings.

Activity Worksheets: These worksheets were designed to guide students through structured tasks related to socioscientific issues and argumentation. They provided insights into the students' ability to engage with the content, organize their thoughts, and construct arguments.

Classroom Discussion Videos: Video recordings of classroom discussions were made to capture students' interactions during the argumentation activities. These videos allowed for an in-depth analysis of how students articulated their arguments, responded to peers, and engaged with the socioscientific topics under discussion.

Face-to-Face Interview Recordings: After the implementation of the activities, face-to-face interviews were conducted with selected students. These interviews aimed to gain a deeper understanding of students' reflections on the activities, their thought processes, and their perspectives on the socioscientific issues discussed. The interviews were audio-recorded to ensure accurate data capture for subsequent analysis.

These data collection tools provided a rich and multifaceted view of the students' engagement with the argumentation process and their development of argumentation skills.

Validity and Reliability

After the student activity sheets were created, they were theoretically grounded in the literature. To evaluate their appropriateness specifically, whether they aligned with Socioscientific Issue-Based Instruction and argumentation practices expert opinions were sought. One expert was a science education specialist, while the second was a science teacher with a master's degree in SSI and expertise in science education. Based on their feedback, the theoretical soundness of the activities and their suitability for 8th-grade students were assessed separately. Necessary revisions were made in accordance with their suggestions. To enhance the validity and reliability of the qualitative data, multiple strategies were applied. First, the triangulation strategy was implemented. Triangulation involves using data obtained through different methods such as observations, interviews, document analysis, video, and audio recordings to cross-validate findings (Yıldırım & Şimşek, 2018). In the second phase, participant confirmation was conducted to further ensure validity and reliability. This process involved presenting research findings back to the participants to prevent misinterpretation of the data (Yıldırım & Şimşek, 2018). Additionally, face-to-face interviews were carried out with students selected through outlier sampling, allowing for a deeper exploration of varying perspectives. This approach strengthened the data triangulation process by complementing written arguments with oral argumentation data. Furthermore, the qualitative data analysis process was described in detail using thick description, ensuring transparency and rigor in interpretation. Lastly, an expert review strategy was applied to reinforce the reliability of the study. In this stage, specialists in qualitative research methods critically examined all aspects of the study including data collection, analysis, and conclusions offering constructive feedback to enhance research quality (Yıldırım & Şimşek, 2018).

Data Analysis

In analyzing the data obtained from student activity sheets, Sadler and Fowler (2006) was used in the argumentation assessment rubric. The argumentation levels assessment rubric consists of five levels. The rubric levels are shown in Table 1.

Table 1. Argumentation Evaluation Rubric Developed by Sadler and Fowler (2006)

Score	Description
0	No Justification
1	Justification with no grounds
2	Justification with simple grounds
3	Justification with elaborated grounds
4	Justification with elaborated grounds and a counterposition

When analyzing students' argumentation levels, not every sentence was accepted as justification. Instead, emphasis was placed on the coherence between their justifications and arguments. Additionally, in the section on persuading the opposing view,

students were not simply attempting to convince their opponents but were primarily engaged in refuting counterarguments, sometimes even challenging their own initial ideas. The analysis was conducted in two different stages by the researcher. During this process, students with varying levels of argumentation skills attempted to convince each other through structured discussions. At the conclusion of the study, students participated in three data collection methods: discussion sessions, written argumentation tasks, and oral interviews. The results were presented in tables and graphical formats, displaying students' argumentation levels, their ability to defend socioscientific issues, and their perspectives based on argumentation performance.

FINDINGS

The study includes a total of eight activities, and reporting on all of them could excessively broaden the scope of the article, potentially hindering the reader's ability to develop a deep understanding of the topic. Additionally, detailed explanations for each activity could disrupt the flow of the article, leading to a shift in focus. In this context, only three activities, genetically modified organisms, nuclear energy, and agricultural pesticides, were selected to allow for a more in-depth analysis and a more effective presentation of the study's findings. This choice begins with addressing the activities at the start of the study, then moves on to more complex and impactful topics, which reflect the essence of the subject. Furthermore, these activities are not only scientifically controversial but also have significant societal implications, making it possible to place these findings within a broader academic context. As a result, focusing solely on these three activities was a strategic decision aimed at enhancing the scientific depth and academic value of the study.

In first activity, students' argumentation skills about GMOs were examined. The students were asked the question "Do you think the genetics of living organisms should be changed?" and students were asked to write down their ideas. The results of written argumentation In light of the data, student levels are shown in Figure 1.

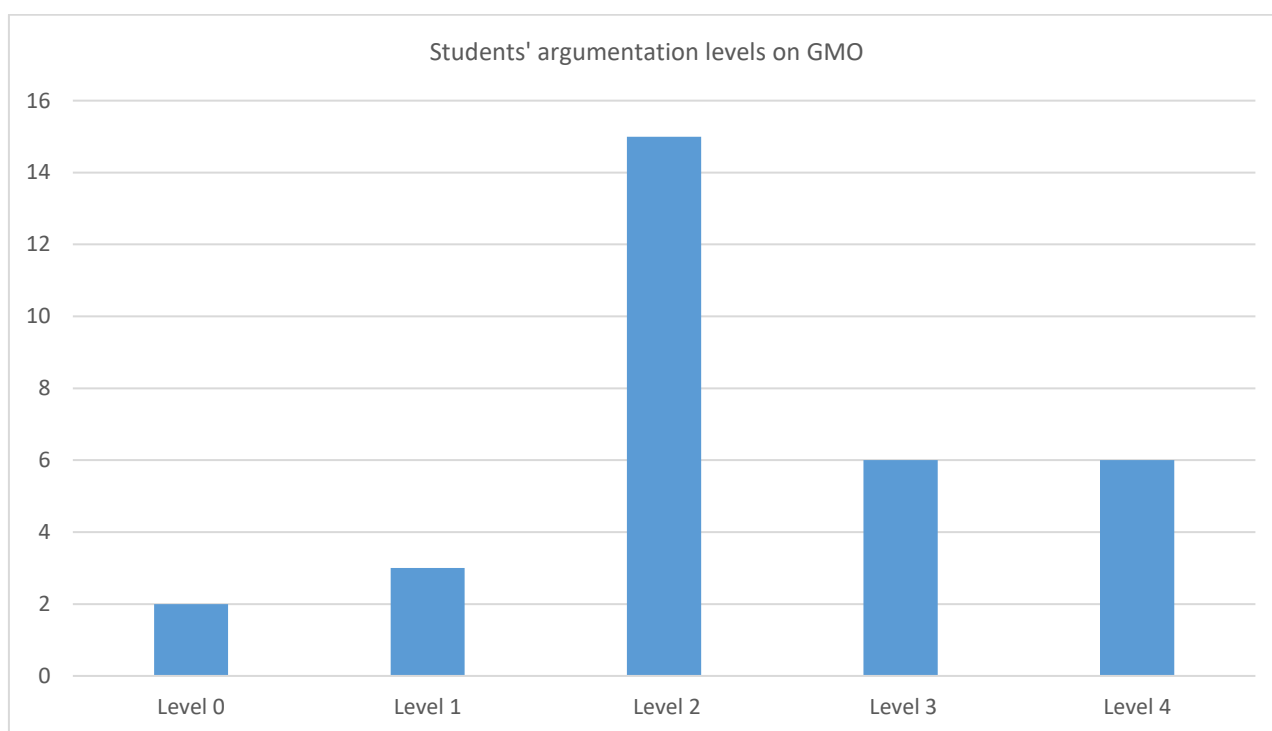


Figure 1. Students' argumentation levels on genetically modified foods

When analyzing the Figure 1, it is observed that 2 students are at Level 0, 3 students at Level 1, 15 students at Level 2, 6 students at Level 3, and 6 students at Level 4. The results indicate that students demonstrated varying levels of argumentation skills. Additionally, their views on GMOs were diverse: 4 students supported GMOs, while 28 opposed them.

Table 2. Arguments For and Against GMOs Based on Preservice Teachers' Responses

Category	Code	Frequency	Excerpt
For GMO	Yield and quality Increase	4	GMOs increase the yield of crops.
	Nutritional need	3	Hunger thanks to GMO crops problems can also be prevented
	Pesticide use reduction	2	Reduces the use of pesticides
	Economy	1	So we can grow more resilient crops our economy grows.
Against GMO	Health	24	GMO products cause cancer
	Ecological balance	20	It also destabilizes the ecosystem
	Biological pollution	8	Genetic pollution occurs when GM pollen is released into the environment
	Economy	3	GMOs will bankrupt farmers

When Table 2 is analyzed, it is evident that most students oppose GMOs. Those in favor of GMOs argue that they enhance crop yield and quality, alleviate food shortages, offer economic benefits, and reduce pesticide use. Conversely, students against GMOs believe they pose health risks, disrupt ecological balance, contribute to biological pollution, and threaten biodiversity. Table 4.26 below presents students' views on GMOs based on their argumentation levels.

Table 3. Distribution of Student Views on GMOs Across Reasoning Levels

Student views	Level 0	Level 1	Level 2	Level 3	Level 4
Increasing yield and quality		1	1	1	1
Nutritional needs			1	1	1
Reducing pesticide use			2		
Economy				1	
Health	2		14	5	3
Ecological balance	2		12	4	2
Biological pollution			3	2	3
Death of living organisms	1		3	1	1
Economy			2	1	

Table 3 presents the distribution of student views on genetically modified organisms (GMOs) across five reasoning levels (Level 0 to Level 4). The thematic categories encompass both supportive and critical perspectives. The majority of higher-level reasoning responses are concentrated in the "health" and "ecological balance" categories, indicating a deeper engagement with the potential risks of GMOs. Conversely, themes such as "reducing pesticide use," "nutritional needs," and "increasing yield and quality" are represented by fewer students and are generally associated with lower or mid-level reasoning. These findings suggest that students are more likely to provide detailed and complex reasoning when discussing perceived negative consequences of GMOs. At the conclusion of the evaluation, interviews were conducted with students representing high, medium, and low argumentation levels. The following excerpt is from an interview with a student at the medium argumentation level:

Researcher: Do you think there are people who hold different views from yours?

Participant: Of course, there are.

Researcher: How do you think they perceive the issue?

Participant: They likely believe that GMOs contain harmful chemicals, making them more dangerous for human health. That's their perspective.

Researcher: How would you persuade those who disagree with you by challenging their arguments?

Participant: If we use GMOs, we can produce higher yields, increase economic gains, and protect crops from pests. Without

GMOs, pests could destroy nearly half of our crops, leading to lower yields and financial losses. Eventually, this could result in food shortages. Countries might even engage in conflicts to secure food supplies, and societal tensions could rise. To prevent such consequences, we should embrace GMOs.

Researcher: Thank you for sharing your thoughts.

In this activity, students' argumentation skills regarding agricultural pesticides were examined. They were asked the question, *"Do you think agricultural pesticides should be used?"* and were requested to write down their opinions. Based on the data obtained from their written argumentation, students' argumentation levels are presented in Figure 2.

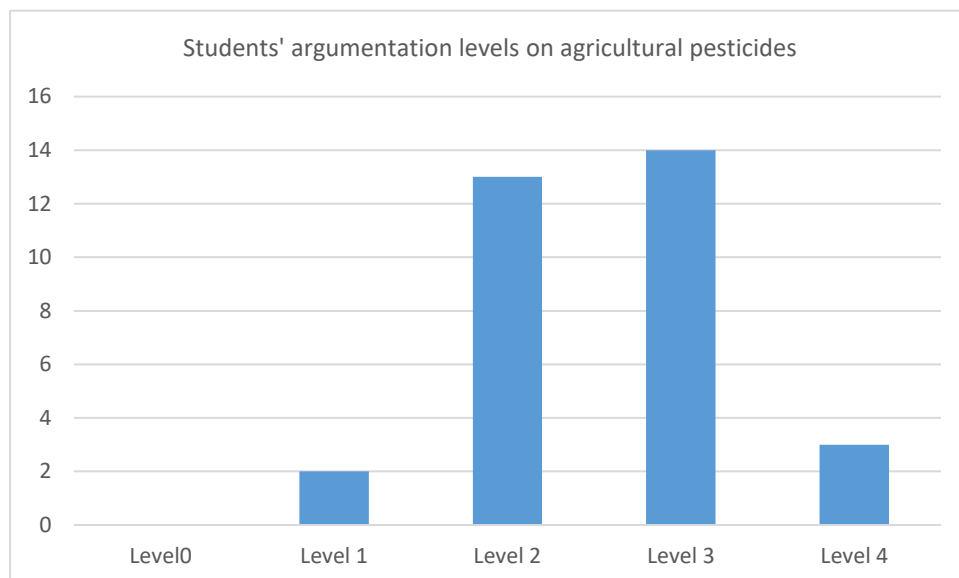


Figure 2. Students' argumentation levels on agricultural pesticides

Upon examining the figure, it is evident that no students were at level 0. Two students demonstrated level 1 argumentation, 13 students reached level 2, 14 students achieved level 3, and 3 students demonstrated level 4 argumentation. Students' stances on the establishment of base stations are presented in Figure 2. It is observed that 19 students support the use of agricultural pesticides, while 13 students oppose it. The number of students who support the establishment of base stations and those who oppose it are very close to each other. The students' views on the use of agricultural pesticides are shown in Table 4.

Table 4 Arguments For and Against the Use of Agricultural Pesticides Based on Student Responses

Category	Code	Frequency	Sample Quote
Agricultural Pesticides Should Be Used	Protects Plants	15	"It protects the plants from harmful insects."
	Increases Yield	18	"Thanks to these pesticides, the yield of our products increases."
	Economy	9	"Without agricultural pesticides, our crops wouldn't grow, and we would face financial difficulties."
	Prevents Famine	8	"Since it increases the productivity of trees, more people can be fed."
	Prevents Trees from Drying	10	"Our trees don't dry out thanks to these pesticides."
Agricultural Pesticides Should Not Be Used	Health	12	"Agricultural pesticides cause cancer."
	Harm to Living Creatures	5	"If we spray pesticides, we also kill harmless creatures."
	Ecological Balance	5	"If we use pesticides, all insects die, and the ecological balance is disrupted."
	Environmental Pollution	4	"It causes soil and environmental pollution."

Upon reviewing Table 4, it becomes evident that students hold diverse perspectives on the use of agricultural pesticides. Those in favor of their use argue that these pesticides protect plants, enhance crop yields, prevent famine, and safeguard trees from drying out. Conversely, students who oppose the use of agricultural pesticides emphasize their detrimental effects on health, the

death of living organisms, the disruption of ecological balance, and environmental pollution. Table 5, presented below, illustrates Distribution of Student Views on Agricultural Pesticides Across Reasoning Levels.

Table 5. Distribution of Student Views on Agricultural Pesticides Across Reasoning Levels

Student Views	Level 0	Level 1	Level 2	Level 3	Level 4
Protects Plants	1	5	7	2	
Increases Yield	1	6	9	2	
Economy	3	5	1		
Prevents Famine	1	3	3	1	
Prevents Trees from Drying	3	5	2		
Health	1	6	4	1	
Harm to Living Creatures	1	2	2		
Ecological Balance	2	2	1		
Environmental Pollution	1	1	1		

Table 5 illustrates the distribution of student responses regarding the use of agricultural pesticides across five levels of reasoning (Level 0 to Level 4). The thematic categories reflect both supportive and opposing views, with frequencies indicating how many students expressed each view at the respective reasoning level. The data reveal that justifications related to "increasing yield" and "protecting plants" are more commonly associated with higher levels of reasoning, suggesting a more developed understanding. In contrast, arguments concerning "environmental pollution" and "ecological balance" tend to be concentrated at lower reasoning levels, indicating less elaboration and conceptual depth in those areas. At the end of the evaluation, interviews were conducted with students who had the highest, medium, and lowest argumentation levels. A part of the interview with a student who demonstrated medium-level argumentation is as follows:

Researcher: So, should agricultural pesticides be used, in your opinion?

Participant: Yes, they should be used.

Researcher: What data do you base your thoughts on?

Participant: They increase the yield significantly. One becomes ten, and ten becomes fifty. It maximizes the yield, and it's already been proven.

Researcher: What are the reasons behind your way of thinking?

Participant: Since they increase the yield, we earn more money, and it's better financially.

In this activity, students' argumentation skills regarding nuclear energy were examined. The question "Do you think nuclear power plants should be built?" was asked to the students, and they were asked to write down their opinions. Based on the written argumentation, the students' argumentation levels are shown in Figure 3.

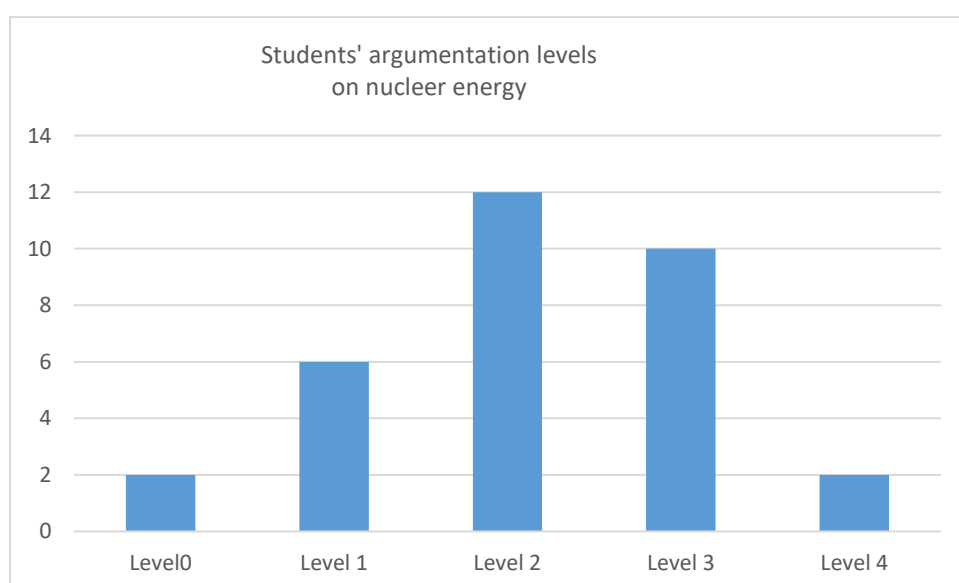


Figure 3. Students' argumentation levels on nuclear energy

When examining Figure 3, it can be seen that 2 students demonstrated level 0 argumentation, 6 students demonstrated level 1, 12 students demonstrated level 2, 10 students demonstrated level 3, and 2 students demonstrated level 4 argumentation. The students' views on the establishment of nuclear power plants are shown in Table 6.

Table 5. Student Arguments For and Against the Construction of Nuclear Power Plants

Category	Code	Frequency	Sample Quote
Nuclear Power Plants Should Be Built	Increased Development	7	"Thanks to nuclear energy, we could become one of the more developed countries in the world."
	High Energy Production	14	"It produces more energy than wind turbines and coal."
	Minimal Environmental Impact	13	"Nuclear energy has very low greenhouse gas emissions."
	Safety	7	"Nuclear power plants are as reliable as airplanes."
	Independent Energy Production	2	"Production can be maintained at the same capacity throughout the year."
	Economic	3	"Nuclear energy would also improve our economy because we wouldn't need to buy energy."
Nuclear Power Plants Should Not Be Built	Risk of Explosion	9	"If nuclear power plants explode, millions of people could die."
	Weapon Production	3	"Weapons that kill people are also produced in these plants."
	Harm to Living Creatures	10	"The radioactive materials it emits cause cancer in people."
	Radiation Emission	6	"The radiation emitted by nuclear energy would make the world uninhabitable."
	Birth Defects and Deaths	6	"Nuclear energy increases the chances of children being born dead or with disabilities."

When examining Table 5, it can be seen that students have very different views on nuclear power plants. Students who support nuclear power plants argue that the establishment of these plants would increase the country's level of development, allow for high levels of energy production, enable electricity generation independent of climate conditions, positively impact the economy, and be reliable. On the other hand, students who oppose nuclear power plants believe that these plants pose a risk of explosion, harm living creatures, emit radiation, could lead to birth defects and stillbirths, and could be used for weapon production. Table 6 below shows Distribution of Student Views on Nuclear Power Plants Across Reasoning Levels

Table 6. Distribution of Student Views on Nuclear Power Plants Across Reasoning Levels

Student Views	Level 0	Level 1	Level 2	Level 3	Level 4
Increase in Development	3	3	1		
High Energy Production	3	5	5	1	
Minimal Environmental Harm	3	5	3	1	
Safety	1	3	2	1	
Ability to Produce Energy Independent of Climate	1	1			
Economic	2	1			
Explosion Risk	1	5	2	1	
Weapon Production	2	1			
Harm to Living Creatures	2	4	3	1	
Radiation Emission	1	3	2		
Birth Defects and Stillbirths	1	2	1	1	

Table 6 displays the distribution of student perspectives on nuclear power plants across five levels of reasoning. The categories include both positive and negative views. Themes such as "high energy production" and "explosion risk" are prominent across multiple reasoning levels, indicating varied depth in students' understanding. Positive aspects like "minimal environmental harm" and "safety" generally appear at mid to higher reasoning levels, while concerns such as "weapon production" tend to cluster at lower levels. This distribution suggests that students engage with both the benefits and risks of nuclear energy, though the complexity of their reasoning varies by theme. At the conclusion of the evaluation, interviews were conducted with students

representing the highest, medium, and lowest levels of argumentation. Below is an excerpt from the interview with a student demonstrating high-level argumentation:

Researcher: Do you think nuclear energy should be established?

Participant: Yes, I believe nuclear energy should be established.

Researcher: And what evidence supports your opinion?

Participant: Nuclear energy should be established because, compared to other power plants, it can generate electricity many times more efficiently and can provide energy year-round.

Researcher: What are the reasons behind your perspective?

Participant: The risk of explosion in nuclear power plants is very low, about one in a thousand. While they do emit radioactive material, it is not harmful. This does not cause significant problems, and in the long term, establishing nuclear plants would bring us closer to other developed countries in terms of advancement. Additionally, it would reduce our dependency on foreign energy.

DISCUSSION, CONCLUSION AND SUGGESTIONS

This study aims to investigate the impact of argumentation-based activities on 8th-grade students' views on socioscientific issues. To achieve this objective, students' perspectives on socioscientific topics were examined following the implementation. Additionally, students' opinions on discussing current socioscientific issues using the argumentation method were explored. Based on these findings, recommendations were provided to guide researchers planning similar activities in future studies.

The students' argumentation skills regarding GMOs were examined, and they were asked, "Do you think the genetics of living organisms should be altered?" It was found that most students were against GMOs. Students supporting GMOs argued that GMOs increase the yield and resilience of crops and prevent food shortages. However, it was observed that these students were fewer in number and less effective in the argumentation process compared to those who opposed GMOs. Literature supports these findings. Bilen and Özel (2012) found that middle school students supported GMOs because they believed it would allow farmers to produce more. Similarly, Topaloğlu (2019) found that seventh-grade students were more likely to support GMOs due to the positive impacts they believed GMOs have on humans, animals, and plants. Students against GMOs argued that they are harmful to human health, disrupt the ecological balance, cause biological pollution, and lead to the death of organisms. Analyzing the students' responses, it was evident that they opposed GMOs primarily due to their perceived harmful effects on health and the ecological balance. Supporting literature, such as the study by Yılmaz, Üner, and Ercan (2015), found that 83.2% of university students believed GMOs are harmful to health, and 78.5% thought GMOs disrupt the ecological balance. Additionally, Çankaya and İşçen (2015) found that almost all science teacher candidates considered GMOs harmful. In this activity, students who opposed GMOs were more successful in generating arguments and refuting the opposing side's points. However, due to students' limited exposure to GMOs and the remote learning setting during their eighth-grade year, many struggled to generate strong arguments. Several students repeated the same arguments raised by others, which were also reflected in their written arguments. As a result, approximately 20 students produced arguments at level 2 or below, highlighting how foundational knowledge plays a crucial role in debates.

The students' argumentation skills regarding agricultural chemicals were also assessed by asking, "Do you think agricultural chemicals should be used?" Nineteen students supported the use of agricultural chemicals, while thirteen opposed them. All students participated in the discussion, expressing their opinions. Many students' families are engaged in farming, and these chemicals are commonly used, which may explain their higher level of interest in the topic. It can be concluded that students are more engaged with socioscientific issues they encounter in daily life, and such topics can lead to more effective learning by fostering active participation. Students supporting the use of agricultural chemicals argued that they protect plants, enhance their resistance, increase yield, prevent tree death, boost economic income, and prevent shortages. Research by Gültekin (2019) aligns with this view, finding that farmers use agricultural chemicals to address plant diseases, harmful insects, and weeds, thereby improving the yield and quality of their crops. Similarly, Uzundumlu, Kılıç, and Tozlu (2017) found that farmers in hazelnut production used agricultural chemicals to eliminate pests, including the hazelnut worm. Students opposing agricultural chemicals argued that these substances harm health, cause deaths, contribute to environmental pollution, and disrupt ecological balance. This view is supported by Özkan, Akçaöz, and Karadeniz (2003), who found that 70.4% of citrus producers in Antalya believed agricultural chemicals left residues in products, and 96.8% of them noted the negative environmental impact of excessive pesticide use. During the discussion, students who actively used these chemicals were particularly vocal, developing diverse arguments to persuade the opposing side. Only two students produced arguments below level 2, indicating their limited knowledge and interest in the topic. Socioscientific issues often involve dilemmas and do not have clear-cut answers (Sadler, 2004). As a result, individuals must make choices when faced with such topics. Even though students recognized the potential harms of agricultural chemicals, some argued that they were necessary, reflecting the complex nature of socioscientific issues. Additionally, some students pointed out that the negative impacts of agricultural chemicals were unethical or even sinful from a religious perspective. This illustrates that engaging with socioscientific topics contributes to the development of students' moral and ethical understanding. Teaching

such issues through argumentation is crucial for nurturing individuals who can think critically, evaluate different perspectives, and contribute to solving real-world problems (Driver, Newton, & Osborne, 2000).

The students' argumentation skills regarding nuclear energy were also explored with the question, "Do you think nuclear power plants should be established?" During the discussion, it was observed that students did not show as much interest in nuclear energy power plants as they did in geothermal energy systems, and fewer students actively participated. This is likely because nuclear power plants are not present in their region, and the students lacked sufficient knowledge about them. While 20 students supported nuclear power plants, 12 were opposed. Those in favor argued that nuclear power plants produce large amounts of electricity, cause less environmental damage, are reliable, generate energy independent of climate conditions, are cost-effective, and contribute to the development of the country. These students highlighted the importance of energy production and self-sufficiency for national development, reflecting an awareness of the challenges faced by a country with growing energy demands and limited resources (Çalışkan, 2009). Furthermore, the fact that most students emphasized the lower environmental impact of nuclear energy indicates that they care about environmental sustainability. Research by Sağlam (2016) found that most teacher candidates supported the construction of nuclear power plants, aligning with the students' views in this study. Similarly, Tekgöz and Yalman (2020) found that science teachers supported nuclear power plants due to their potential to meet energy needs and reduce dependence on foreign sources. Opponents of nuclear energy argued that nuclear power plants pose a risk of explosions, release harmful radiation, harm living organisms, and cause stillbirths. These students were particularly concerned about the health and safety risks associated with nuclear power plants, suggesting that they valued human health over energy production. Sürmeli, Duru, and Duru (2017) found similar concerns among teachers regarding the risks of nuclear accidents and environmental degradation caused by nuclear energy. Although 130 students made efforts to defend their positions, the lack of participation from some students prevented the discussion from reaching the desired level of argumentation. The fact that students had never seen a nuclear power plant or fully understood how they generate energy contributed to their inability to produce advanced arguments. Therefore, it is important to select socioscientific issues that students are likely to encounter in their daily lives, as they are more likely to engage deeply with these topics. Furthermore, it is essential to provide foundational knowledge before debates to ensure that students can participate in meaningful discussions.

This study aimed to assess students' argumentation skills regarding genetically modified organisms, agricultural pesticides, and nuclear energy. The findings revealed that most students expressed opposition to GMOs, emphasizing health risks and ecological imbalance. However, those in favor of GMOs presented arguments related to increased crop yield, resilience, and food security. Similarly, in the debate about agricultural pesticides, students who supported their use highlighted benefits such as pest control and increased agricultural productivity, while opponents raised concerns about health risks and environmental harm. In the discussion on nuclear energy, while some students supported it due to its potential for high energy production and environmental benefits, others expressed concerns about safety risks and the potential for harm to living organisms.

The results underscore the importance of students' knowledge in shaping their arguments, as well as the influence of real-life experiences with socioscientific issues on their engagement and perspectives. It also became evident that while students demonstrated an understanding of the issues, many struggled with higher-level argumentation due to limited exposure to relevant information and limited opportunities for in-depth discussions.

It is recommended that researchers enhance informational support for future studies. To improve argumentation skills, students need a solid foundation of knowledge on complex socioscientific issues. This can be achieved by incorporating more detailed lessons, case studies, and debates that present diverse perspectives. On the other hand, Teachers should cultivate an environment that promotes active student engagement, particularly in debates on controversial topics. Encouraging all students to express their views, critically evaluate opposing opinions, and construct well-supported arguments can significantly enhance their critical thinking and argumentation skills. Further, effective facilitation of complex debates requires specific skills. Teachers should receive training to guide students in forming coherent arguments, critically analyzing opposing viewpoints, and using evidence to substantiate their claims. Such training could be integrated into ongoing professional development programs for science educators.

By implementing these recommendations, educators can better prepare students to engage thoughtfully with socioscientific issues, enhancing their understanding of complex problems and their ability to argue and debate effectively.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, author-ship, and/or publication of this article.

Statements of publication ethics

We hereby declare that the study has not unethical issues and that research and publication ethics have been observed carefully.

Author Contribution Statement:

The research process's theoretical framework, methodology, and design were carried out in collaboration with the first and second authors. Data collection was carried out by the second author. Data analysis, evaluation of the findings was carried out by both authors, and manuscript drafting were carried out by the first author. The first author contributed to revising the study's methodological approach and the manuscript's final editing. Both authors read and approved the final version of the article.

Researchers' contribution rate

The study was conducted with equal collaboration of the researchers however it was reported by the first author.

Ethics Committee Approval Information

Ethics committee approval for the study was obtained at the Aydın Adnan Menderes University Ethics Commission meeting dated 17.11.2020 and numbered E.59911 with the research code 2020/18-III.

REFERENCES

- Alkın Şahin, S., Tunca, N., & Ulubey, Ö. (2014). The relationship between pre-service teachers' educational beliefs and their critical thinking tendencies. *İlköğretim Online*, 13(4), 1473–1492.
- Babacan, M. (2017). *Sosyobilimsel konulardaki etkinliklerin yedinci sınıf öğrencilerinin eleştirel düşünme becerilerine etkisi* [Yayınlanmamış yüksek lisans tezi]. Niğde Üniversitesi.
- Başar, S. (2018). *Fen bilimleri öğretmen adaylarının fende matematiğin kullanımına yönelik özyeterlik inançları, 21. yy becerileri ve aralarındaki ilişkinin incelenmesi* [Yayınlanmamış yüksek lisans tezi]. Hacettepe Üniversitesi.
- Bench-Capon, T. (2003). Persuasion in Practical Argument Using Value-based Argumentation Frameworks. *J. Log. Comput.*, 13, 429–448. <https://doi.org/10.1093/logcom/13.3.429>.
- Bilen, K., & Özel, M. (2012). Üstün yetenekli öğrencilerin biyoteknolojiye yönelik bilgileri ve tutumları. *NEF-EFMED*, 6(2), 135–152.
- Birdal, H. (2019). *Sosyobilimsel konularda argümantasyona dayalı öğrenme uygulamalarının fen bilimleri öğretmen adaylarının öğrenciyi anlama bilgilerinin gelişimine etkisi* [Yayınlanmamış yüksek lisans tezi]. Kafkas Üniversitesi.
- Cleisey, M., & Kazibwe, F. (2025). Assessing knowledge and understanding of genetically modified organisms (GMOs) among university students in Uganda. *East African Journal of Interdisciplinary Studies*. <https://doi.org/10.37284/eajis.8.1.2953>
- Creswell, J. W. (2012). *Educational research: Planning, conducting, and evaluating quantitative and qualitative research* (4th ed.). Pearson Education.
- Çalık, D., & Çınar, Ö. P. (2009). Geçmişten günümüze bilgi yaklaşımları bilgi toplumu ve internet. *XIV. Türkiye'de İnternet Konferansı Bildirileri*. Bilgi Üniversitesi.
- Çalışkan, Ş. (2009). Türkiye'nin enerjide dışa bağımlılık ve enerji arz güvenliği sorunu. *Dumlupınar Üniversitesi Sosyal Bilimler Dergisi*, 25, 297–310.
- Çankaya, C., & İşçen, F. C. (2015). Fen bilgisi öğretmen adaylarının genetiği değiştirilmiş organizmalara (GDO) dair bilgi düzeylerinin ve görüşlerinin belirlenmesi. *International Journal of Social Science*, 32, 537–554.
- Dawson, V. (2025). Using socioscientific issues to teach argumentation to Year 7 science students in a low socioeconomic rural Australian school. *Research in Science Education*. <https://doi.org/10.1007/s11165-024-10224-y>
- Dub, A. (2023). The impact of innovations and technological development on modern society and global dynamics. *Economic Affairs*. <https://doi.org/10.46852/0424-2513.4.2023.39>
- Driver, R., Newton, P., & Osborne, J. (2000). Establishing the norms of scientific argumentation in classrooms. *Science Education*, 84(3), 287–312.
- Durant, J. (1993). What is scientific literacy. In J. R. Durant & J. Gregory (Eds.), *Science and culture in Europe* (pp. 129–137). Science Museum.
- Eastwood, J. L., Sadler, T. D., Zeidler, D. L., Lewis, A., Amiri, L., & Applebaum, S. (2012). Contextualizing nature of science instruction in socioscientific issues. *Journal of Research in Science Teaching*, 49(12), 2289–2315.
- Golovach, L. (2023). Self-training of future skilled workers using innovative learning technologies in the educational process. *Ukrainian Professional Education*. <https://doi.org/10.33989/2519-8254.2023.14.300701>
- Gültekin, T. (2019). *Kırşehir ili Mucur ilçesi hububat üreticilerinin bitki koruma yönünden karşılaştıkları sorunlar ile tarımsal ilaç kullanım durumunu etkileyen faktörlerin belirlenmesi* [Yayınlanmamış yüksek lisans tezi]. Kırşehir Ahi Evran Üniversitesi.
- Huda, M. (2023). Trust as a key element for quality communication and information management: Insights into developing safe cyber-organisational sustainability. *International Journal of Organizational Analysis*. <https://doi.org/10.1108/ijoa-12-2022-3532>
- Jiang, M. (2024). A review of the impacts of industrial revolutions in world history. *Communications in Humanities Research*. <https://doi.org/10.54254/2753-7064/39/20242245>
- Karaca, H. S. (2018). *Yapılandırmacı yaklaşım yoluyla sosyobilimsel konulara dayalı fen eğitiminin 7. sınıf öğrencileri üzerine etkileri* [Yayınlanmamış yüksek lisans tezi]. Trakya Üniversitesi.
- Karamanlı, E. (2019). *Sosyobilimsel konularda sınıf içi destekli blog uygulamaları ile ortaokul öğrencilerinin argümantasyon düzeylerinin ve informal akıl yürütme örüntülerinin incelenmesi* [Yayınlanmamış yüksek lisans tezi]. Mersin Üniversitesi.
- Karışan, D. (2011). *Fen bilgisi öğretmen adaylarının iklim değişiminin dünyamıza etkileri konusundaki yazılı argümantasyon yeteneklerinin incelenmesi* [Yayınlanmamış yüksek lisans tezi]. Yüzüncü Yıl Üniversitesi.
- Karmana, W. (2024). Penerapan model project based learning (PjBL) terhadap kemampuan literasi sains dan hasil belajar peserta didik pada pembelajaran IPA di sekolah. *Panthera: Jurnal Ilmiah Pendidikan Sains dan Terapan*. <https://doi.org/10.36312/panthera.v4i2.273>

- Kurylo, V., Karaman, O., Bader, S., Pochinkova, M., & Stepanenko, V. (2023). Critical thinking as an information security factor in the modern world. *Social Legal Studios*. <https://doi.org/10.32518/sals3.2023.67>
- Levinson, R. (2006). Towards a theoretical framework for teaching controversial socio-scientific issues. *International Journal of Science Education*, 28(11), 1201–1224.
- Macalalag, A., Kaufmann, A., Van Meter, B., Ricketts, A., Liao, E., & Ialacci, G. (2024). Socioscientific issues: Promoting science teachers' pedagogy on social justice. *Disciplinary and Interdisciplinary Science Education Research*. <https://doi.org/10.1186/s43031-024-00118-4>
- Mammadov, H., & Gulamova, A. (2025). Impact of genetically modified organisms on human health. *Bulletin of Science and Practice*. <https://doi.org/10.33619/2414-2948/111/18>
- Martin, M., Enns, K., & Stewart, K. (2016). New perspectives on teaching GMOs. *The Agricultural Education Magazine*, 88, 19.
- Meral, E. (2018). *Argümantasyon tabanlı bilim öğrenme yaklaşımının öğrencilerin akademik başarılarına, eleştirel düşünme eğilimlerine ve argüman oluşturma becerilerine etkisi* [Yayınlanmamış yüksek lisans tezi]. Atatürk Üniversitesi.
- Ministry of National Education (MoNE) Milli Eğitim Bakanlığı MEB. (2013). *İlköğretim kurumları fen bilimleri dersi öğretim programı*. Ankara.
- Özdemir, O. (2010). Fen ve teknoloji öğretmen adaylarının fen okuryazarlığının durumu. *Türk Fen Eğitimi Dergisi*, 7(1), 42–56.
- Özkan, B., Akçağöz, H. V., & Karadeniz, C. F. (2003). Antalya ilinde turuncgil üretiminde tarımsal ilaç kullanımına yönelik üretici tutum ve davranışları. *Anadolu Dergisi*, 13(2), 103–116.
- Patronis, T., Potari, D., & Spiliotopoulou, V. (1999). Öğrencilerin sosyo-bilimsel bir konuda karar alma sürecindeki argümantasyonu: Öğretim için çıkarımlar. *Uluslararası Bilim Eğitimi Dergisi*, 21(7), 745–754.
- Prävălie, R., & Bandoc, G. (2018). Nuclear energy: Between global electricity demand, worldwide decarbonisation imperativeness, and planetary environmental implications. *Journal of Environmental Management*, 209, 81–92. <https://doi.org/10.1016/j.jenvman.2017.12.043>
- Sadler, T. D. (2004). Informal reasoning regarding socioscientific issues: A critical review of research. *Journal of Research in Science Teaching*, 41(5), 513–536.
- Sadler, T. D., & Fowler, S. R. (2006). A threshold model of content knowledge transfer for socioscientific argumentation. *Science Education*, 90(6), 986–1004.
- Sadler, T. D., & Zeidler, D. L. (2004). The morality of socioscientific issues: Construal and resolution of genetic engineering dilemmas. *Science Education*, 88(1), 4–27.
- Sadler, T. D., & Zeidler, D. L. (2005). The significance of content knowledge for informal reasoning regarding socioscientific issues: Applying genetics knowledge to genetic engineering issues. *Science Education*, 89(1), 71–93.
- Sağlam, H. G. (2016). *Öğretmen adaylarının nükleer enerji kullanımına yönelik informal muhakemeleri üzerine karma yöntem araştırması* [Yayınlanmamış yüksek lisans tezi]. Aksaray Üniversitesi.
- Sevgi, Y. (2016). *Gazete haberlerindeki sosyobilimsel konuların argümantasyon yöntemiyle tartışılmasının ortaokul 7. sınıf öğrencilerinin eleştirel düşünme, karar verme ve argümantasyon becerilerine etkisi* [Yayınlanmamış yüksek lisans tezi]. Marmara Üniversitesi.
- Smit, R., Rietz, F., & Büchel, D. (2025). Using the socioscientific issue approach to foster secondary students' argumentation skills, science self-efficacy beliefs and science interest. *International Journal of Science Education*. <https://doi.org/10.1080/09500693.2025.2460050>
- Sürmeli, H., Duru, N., & Duru, R. (2017). Nükleer enerji ve nükleer santraller konusuna yönelik öğretmen tutumlarının farklı değişkenler açısından incelenmesi. *Necatibey Eğitim Fakültesi Elektronik Fen ve Matematik Eğitimi Dergisi (EFMED)*, 11(1), 293–319.
- Tekgöz, T., & Ercan-Yalman, F. (2020). Nükleer santraller hakkında fen bilgisi öğretmenlerinin görüşü: Akkuyu örneği. *Muğla Sıtkı Koçman Üniversitesi Eğitim Fakültesi Dergisi [MSKU Journal of Education]*, 7(2), 144–158.
- Topaloğlu, Y. M. (2019). Okul dışı öğrenme ortamının öğrencilerin GDO ile ilgili görüşlerine etkisinin incelenmesi. *Anadolu Öğretmen Dergisi*, 3(2), 160–183.
- Topçu, M. S. (2015). *Sosyobilimsel konular ve öğretimi*. Pegem Akademi Yayıncılık.
- Toklu, O., Tunçer, F., & Vardar, N. (2025). Environmental impact assessment of nuclear power plants. *International Journal of New Findings in Engineering, Science and Technology*. <https://doi.org/10.61150/ijonfest.2025030107>
- Türköz, G. (2019). Fen bilgisi öğretmen adaylarının çeşitli sosyo-bilimsel konulara yönelik kararlarının, gerekçelerinin ve argüman kalitelerinin incelenmesi: Youtube destekli sınıf içi tartışma kullanımı. *Sinop Üniversitesi Fen Bilimleri Enstitüsü Dergisi*.
- Uzundumlu, A. S., Kılıç, B., & Tozlu, G. (2017). Fındık üretiminde kimyasal ilaç kullanımını etkileyen faktörlerin analizi: Giresun ili örneği. *Gümüşhane Üniversitesi Fen Bilimleri Enstitüsü Dergisi*, 7(1), 1–9.
- Viehmann, C., Cárdenas, J., & Peña, C. (2024). The use of socioscientific issues in science lessons: A scoping review. *Sustainability*, 16(14), Article 5827. <https://doi.org/10.3390/su16145827>
- Yıldırım, A., & Şimşek, H. (2018). *Sosyal bilimlerde nitel araştırma yöntemleri* (12th ed.). Seçkin Yayıncılık.
- Yılmaz, B., Üner, A. K., & Ercan, A. (2015). Üniversite öğrencilerinin biyoteknoloji ve genetiği değiştirilmiş gıdalar ile ilgili tutumları. *Akademik Gastroenteroloji Dergisi*, 14(2), 64–71.
- Yin, R. K. (2017). *Case study research and applications: Design and methods* (6th ed.). Sage Publications.
- Zeidler, D. L., & Nichols, B. H. (2009). Socioscientific issues: Theory and practice. *Journal of Elementary Science Education*, 21(2), 49–58.
- Xiao, S. (2020). Rhetorical use of inscriptions in students' written arguments about socioscientific issues. *Research in Science Education*, 50(1), 1–17. <https://doi.org/10.1007/s11165-018-9730-y>



| Research Article / Araştırma Makalesi |

Examining the Opinions of Expert Teachers on the Application of Teaching Career Stage within the Framework of General Competencies of Teaching Profession

Uzman Öğretmenlerin Öğretmenlik Kariyer Basamağı Uygulamasına İlişkin Görüşlerinin Öğretmenlik Mesleği Genel Yeterlikleri Çerçevesinde İncelenmesi¹

Sevgi YENER², Süleyman Nihat ŞAD³

Keywords

1. Teaching Career Steps
2. Expert Teaching Career Steps
3. Expert Teaching
4. General Competencies Of The Teaching Profession

Anahtar Kelimeler

1. Öğretmenlik Kariyer Basamakları
2. Uzman Öğretmenlik Kariyer Basamağı
3. Uzman Öğretmenlik
4. Öğretmenlik Mesleği Genel Yeterlikleri

Received/Başvuru Tarihi
03.05.2025

Accepted / Kabul Tarihi
10.10.2025

Abstract

Purpose: This research aims to examine the application of the teaching career ladder within the framework of General Competencies for the Teaching Profession; in terms of the quality of the trainings received before the exam, the characteristics measured by the exam and its contribution to professional development, in line with the opinions of expert teachers.

Design/Methodology/Approach: A total of 550 teachers who took the expert teacher exam in Konya Selçuklu district and received the title of expert teacher participated in the research, selected through cluster and criterion sampling for the quantitative phase and maximum diversity sampling for the qualitative phase. In the research designed according to the convergent parallel mixed design, quantitative data were collected with 3 scales developed by the researcher, and qualitative data were collected with a semi-structured interview form. Quantitative data were analyzed with Mann Whitney U, Kruskal Wallis, One-Way ANOVA and Spearman Correlation analyses using the SPSS 25 program; qualitative data were analyzed with NVivo 12 Plus.

Findings: While a significant difference was found between the gender variable and teacher opinions in the quantitative findings, no significant relationship was found between the branch variable and teacher opinions. There was no significant relationship between the exam score and seniority variables and teacher opinions. In the qualitative findings, themes related to the effects of education content, exams and titles were determined.

Highlights: It was concluded that the application of teaching career steps did not meet the expectations on professional competencies and that a more objective, performance-based evaluation system was needed.

Öz

Çalışmanın amacı: Bu araştırma, Öğretmenlik Mesleği Genel Yeterlikleri çerçevesinde öğretmenlik kariyer basamağı uygulamasını; sınav öncesinde alınan eğitimlerin niteliği, yapılan sınavın ölçtüğü özellikler ve mesleki gelişime katkısı bakımından uzman öğretmenlerin görüşleri doğrultusunda incelemeyi amaçlamaktadır.

Materyal ve Yöntem: Araştırmaya Konya Selçuklu ilçesinde uzman öğretmenlik sınavına girerek uzman öğretmen unvanı alan öğretmenler arasından nicel aşama için küme ve ölçüt örnekleme, nitel aşama için maksimum çeşitlilik örnekleme ile seçilen toplam 550 öğretmen katılmıştır. Yakınsayan paralel karma desene göre tasarlanan çalışmada nicel veriler araştırmacı tarafından geliştirilen 3 ölçekle, nitel veriler ise yarı yapılandırılmış görüşme formuyla toplanmıştır. Nicel veriler SPSS 25 programı kullanılarak Mann Whitney U, Kruskal Wallis, Tek Yönlü ANOVA ve Spearman Korelasyon analizleriyle; nitel veriler ise NVivo 12 Plus ile analiz edilmiştir.

Bulgular: Nicel bulgulara cinsiyet değişkeni ile öğretmen görüşleri arasında anlamlı fark bulunurken, branş değişkeni ile bulunamamıştır. Sınav puanı ve kıdem değişkenleri ile öğretmen görüşleri arasında anlamlı ilişki yoktur. Nitel bulgulara eğitim içerikleri, sınav ve unvanın etkilerine ilişkin temalar belirlenmiştir.

Önemli Vurgular: Öğretmenlik kariyer basamakları uygulamasının mesleki yeterlikler üzerine beklentileri karşılamadığı ve daha objektif, performansa dayalı bir değerlendirme sistemine ihtiyaç duyulduğu sonucuna varılmıştır.

¹ This study is derived from Sevgi YENER's master's thesis entitled "Uzman Öğretmenlerin Öğretmenlik Kariyer Basamağı Uygulamasına İlişkin Görüşlerinin Öğretmenlik Mesleği Genel Yeterlikleri Çerçevesinde İncelenmesi," conducted under the supervision of Süleyman Nihat ŞAD.

² Konya Selçuklu District Directorate of National Education, 15 July Martyrs Primary School, Konya, Türkiye; <https://orcid.org/0009-0007-5856-3585>

³ Inonu University, Institute of Education Sciences, Department of Curriculum and Instruction, Malatya, Türkiye; <https://orcid.org/0000-0002-3169-2375>

INTRODUCTION

In the globalizing world, competition between countries is increasing. In order to combat this competition, countries need qualified human resources. Therefore, the most important task of education systems is to raise individuals equipped with the knowledge and skills appropriate to the needs of the future. The education system of a nation serves two basic purposes: to raise qualified individuals and to instill active citizenship awareness. In line with these purposes, each education system defines the ideal human model and carries out activities in line with human resources policies and the adopted educational philosophy (Karagözoğlu, 2003).

The teaching profession is gaining a different appearance in the transition from industrial society to information society. In the information society, the production of information is now a priority and the production of objects has given way to the production of information (Doğan, 2016). In the information society, teachers should know their students well, respect each student's individuality, have an unprejudiced attitude and democratize the teaching environment. Therefore, teachers need to constantly improve themselves (Üstüner, 2002).

These changes have made it necessary for teachers to continue their professional development. Rapid changes in information and technology cause the knowledge that teachers have to become obsolete quickly, and this situation causes teachers' competencies to be constantly questioned. For this reason, the need for teachers who can keep up with scientific and technological developments is increasing today and in the future. Our country has also focused on the goal of training qualified manpower by adapting to these global conditions and has adopted the goal of determining teacher competencies (MEB, 2017).

Three different teacher competency certificates have been published in Türkiye to date. Studies to determine teacher competencies began in 1998, the first document was published in 2002 and was divided into three main headings: general culture, special field, and teaching-teaching competencies. There are 6 general headings in the new document published in 2006. These headings were announced as "professional development, getting to know the student, teaching and learning process, monitoring and evaluation of learning, school, family and community relations, program and content knowledge" (MEB, 2006). Eleven years later, a new updated competency certificate was published in 2017. In this document, the competencies are grouped under three general headings: professional knowledge, professional skills, attitudes and values (MEB, 2017). The General Competencies for Teaching Profession, which were renewed in our country, reached their final form after a comprehensive work process, taking into account the teacher standards determined in many countries.

The teaching profession is considered a special profession that undertakes the duties of education, training and management of the state. If teachers choose this profession, they continue to teach until they retire. However, this situation can reduce teachers' motivation and negatively affect their professional development. For this reason, the need to create teacher career ladders has arisen. Research shows that the lack of motivation of teachers in developed and developing countries is a significant problem (Crehan, 2016). However, the issue of creating teacher career ladders has caused discussions on whether it benefits the professional development of teachers (Kaplan and Gülcan, 2020).

In light of all these developments in our country, in order to contribute to the professional development of teachers, in 2005, the Ministry of National Education amended Article 43 of the Basic Law on National Education No. 1739, stipulating that the teaching profession would be divided into three career steps after the candidate period: teacher, expert teacher and head teacher. Later, in accordance with this change in the law, the Regulation on Promotion in Teacher Career Steps was accepted by the Ministry of National Education in 2005 and published in the Official Gazette (MEB, 2005). However, the Regulation on Promotion in Teacher Career Steps prepared after this change in the law was taken to court by the union and some groups, and the Regulation on Promotion in Teacher Career Steps was annulled by the decision of the Constitutional Court. For this reason, although the law on career steps was in force, it could not be implemented (Constitutional Court, 2009).

In 2022, the career stage exam was brought to the agenda again within the scope of the Teaching Profession Law. According to Law No. 7354 published in the Official Gazette on February 14, 2022, teachers who have at least ten years of service in teaching and have completed the necessary professional development studies are entitled to take the exam when they complete the 180-hour Expert Teacher Training Program. Similarly, teachers who have completed ten years in Expert teaching have gained the right to take the exam held on November 19, 2023 when they complete the 240-hour Head Teacher Training Program. Those with a master's degree are exempt from the expert teacher exam, and those with a doctorate are exempt from the head teacher exam, but they are required to attend professional development training (MEB, 2022).

The General Competencies for the Teaching Profession, restructured in 2017, and the Career Steps in Teaching Profession implementation, which came into effect in 2022, are significant recent developments in the advancement of the teaching profession. In the literature, only one study has been found under these two topics, which was conducted solely using qualitative methods and focused on teachers' opinions. This research aims to examine the subject in depth through both qualitative and quantitative dimensions, based on teachers' perspectives, and in relation to variables such as gender, branch, exam scores, and seniority. There appears to be a lack of comprehensive studies in the literature that address the relationship between teaching competencies and the career steps implementation, considering these variables together. Especially, evaluating the quality of pre-exam training, the assessment adequacy of the exam, and the contribution of the career title to professional development within the framework of teachers' views is important for filling this gap in the literature. Accordingly, the research questions developed in this study are directly related to the theoretical background and current practices in the literature.

The main objective of the study is to evaluate the implementation of the teaching career steps based on the views of expert teachers, within the framework of the General Competencies for the Teaching Profession, in terms of the quality of pre-exam training, the features assessed by the exam, and its contribution to professional development.

In line with this main objective, the following research questions are addressed:

1. To what extent do expert teachers think that the training program taken before the teaching career stage exam contributes to the General Competencies of the Teaching Profession?
 - 1a. Do the opinions of expert teachers differ significantly according to their gender?
 - 1b. Do the opinions of expert teachers differ significantly according to their branches?
 - 1c. Is there a significant relationship between expert teachers' opinions and exam scores?
 - 1d. Is there a significant relationship between the opinions of expert teachers and their seniority?
2. To what extent do expert teachers think the Teaching Career Stage Exam can measure the General Competencies of the Teaching Profession?
 - 2a. Do the opinions of expert teachers differ significantly according to their gender?
 - 2b. Do the opinions of expert teachers differ significantly according to their branches?
 - 2c. Is there a significant relationship between the opinions of expert teachers and exam scores?
 - 2d. Is there a significant relationship between the opinions of expert teachers and their seniority?
3. To what extent do expert teachers think that receiving the title of expert teacher contributes to the General Competencies of the Teaching Profession?
 - 3a. Do the opinions of expert teachers differ significantly according to their gender?
 - 3b. Do the opinions of expert teachers differ significantly according to their branches?
 - 3c. Is there a significant relationship between the opinions of expert teachers and exam scores?
 - 3d. Is there a significant relationship between the opinions of expert teachers and their seniority?
4. What are the experiences and opinions of expert teachers regarding the teaching career ladder application?

METHOD/MATERIALS

Research Design

This study, which investigates the teaching career ladder practice reintroduced in 2022 within the framework of the General Competencies for the Teaching Profession last updated in 2017 in our country in terms of the quality of pre-exam training, the competencies assessed by the exam, and its impact on professional development based on the experiences and views of expert teachers, employed the convergent parallel mixed method design, a type of mixed research approach.

In this study, the implementation of the career steps in teaching is considered a multidimensional structure, involving aspects such as the quality of pre-exam training, the competencies assessed through the exam, and its effects on professional development. Therefore, using only a quantitative or only a qualitative method would be insufficient. For this reason, both quantitative data, which reveal numerical trends, and qualitative data, which uncover the reasons behind these trends and the personal experiences of teachers, were evaluated together. According to Creswell and Clark (2018), this approach is particularly useful when the researcher is experienced in both quantitative and qualitative methods, as it is based on the assumption that each data type provides different kinds of information.

The reason for choosing a convergent parallel design is that it allows both types of data to be collected simultaneously and analyzed independently and then interpreted in a complementary manner. In this way, teachers' views on the career steps process were examined holistically, both in the context of general trends among a large participant group and through individual experiences. Parallel mixed sampling methods offer the opportunity to compare or integrate quantitative and qualitative sample types. This approach allows for a comprehensive interpretation by combining numerical data with qualitative insights (Creswell & Clark, 2018).

In this context, the main objective of the study—to reveal the impact of the implementation on professional competencies in both a generalizable and in-depth manner—necessitated the use of a mixed method design. In the research, quantitative and qualitative data were collected simultaneously, analyzed separately, and then integrated in a comparative and complementary way. While the quantitative data revealed teachers' attitudes and perceptions regarding the career steps process, the qualitative data explained the reasons, meanings, and contextual details behind these perceptions. Thus, the findings obtained from different types of data supported each other, providing a more comprehensive evaluation.

Quantitative Stage

Working Group

In the quantitative phase of the study, the study group was selected using cluster sampling and criterion sampling methods. First, schools in the Selçuklu district of Konya were determined by dividing them into clusters according to their grades. In cluster sampling, generally all units within the cluster are included in the sample and diversity is tried to be preserved as much as possible within the clusters. However, when the units within the cluster are very similar to each other, it would be more appropriate to select a sub-sample instead of the entire cluster (Kaya, 2015). For this purpose; Later, using criterion sampling, which is one of the non-random sampling methods, only teachers who passed the teaching career stage exam and were entitled to receive the title of expert teacher were determined within the schools and participants who met this criterion were included in the scope of the study. The sample size was calculated using the formula $n = z^2 \cdot N \cdot p \cdot q / (N \cdot d^2 + z^2 \cdot p \cdot q)$ (Sencan, 1993, as cited in Cındık & Akyüz, 1996). Based on this formula, the sample size was calculated as $n = 1.96^2 \cdot 6095 \cdot 0.50 \cdot 0.50 / (6095 \cdot 0.04^2 + 1.96^2 \cdot 0.50 \cdot 0.50) = 546.435$, which was rounded to 546. A sample size of 546 was accepted, and this number was reached. The number 550 was taken as the upper limit for the sample size, and when this number was reached, the survey was concluded. In the criterion sampling method, situations that meet a set of criteria are examined. The researcher can create the criteria himself or use a previously prepared list of criteria (Marshall and Rossman, 2014). The study group was formed by 550 teachers who were selected from 6095 expert teachers working under the Ministry of National Education in Selçuklu district of Konya province, who participated in the expert teacher training program with the determined sampling methods and qualified to receive expert teacher status by taking the expert teacher career stage exam.

Data Collection Tools

In the quantitative phase of the study, 3 different scales consisting of 5-point Likert-type scale items, each consisting of 11 items, developed by the researchers were used as data collection tools. The scales were developed to examine the opinions of expert teachers on the expert teacher career stage education, exam and title within the framework of the General Competencies of the Teaching Profession. The names of the scales are Expert Teacher Career Stage Education Evaluation Scale, Expert Teacher Career Stage Exam Evaluation Scale, Expert Teacher Title Evaluation Scale. While developing the scales, in order to ensure content validity, expert opinions were received from a philosophy expert faculty member, a measurement and evaluation expert faculty member, a guidance and psychological counseling expert faculty member, a program development expert faculty member and a Turkish language education expert teacher, and the total number of questions was reduced to 33 by making the necessary corrections. The preliminary application of 3 different scales, each consisting of 11 items, was carried out for 250 expert teachers working in different branches in Selçuklu district of Konya province.

In the analysis of the obtained data, Kaiser-Meyer Olkin (KMO) and Bartlett's Test of Sphericity tests were applied to evaluate the suitability of the scales for factor analysis. KMO values were above 0.50 and were accepted as perfect in the range of 0.90–1.00 (Field, 2009). The significance of the Bartlett test also confirms that there are important relationships between the variables in the data set and that it is a suitable structure for factor analysis (Büyüköztürk, 2020, p. 136). As a result of the exploratory factor analysis, it was determined that each scale consisted of a single factor. For the Expert Teacher Career Stage Education Evaluation Scale, the factor loadings ranged between 0.89 and 0.95, and an example item for the scale could be "To what extent did the training program offered before the expert teacher career stage exam contribute to your skills in managing the teaching and learning process?" The factor loadings of the Expert Teacher Career Stage Exam Evaluation Scale range from 0.87 to 0.95, and an example item for this scale can be presented as "To what extent does the expert teacher career stage exam measure your ability to establish effective communication and cooperation with students, colleagues, families and other stakeholders in education?" The factor loadings of the Expert Teacher Title Evaluation Scale range from 0.95 to 0.97, and an example item for this scale can be presented as "To what extent did being an expert teacher contribute to your personal and professional development through self-evaluation?" It was determined that the factors largely explained the total variance in all scales.

According to the results of confirmatory factor analysis conducted with AMOS 22.0, the fit values of the models were found to be consistent with the values given in the table. In addition, the total item correlations of the items were examined and it was seen that the scales were aimed at measuring the contribution of the education, exam and title received before the exam to the General Competencies of the Teaching Profession. Finally, Cronbach Alpha coefficients were calculated in order to determine the reliability of the scales and it was determined that the coefficients were over 0.70. A Cronbach Alpha coefficient over 0.70 is considered sufficient for the reliability of that measurement tool (Büyüköztürk, 2020, p. 183). As a result of the analyzes, it was decided that the scales were applicable to expert teachers, reliable and useful. The data related to the analyzes are presented in Table 1.

Table 1. Results of Reliability Analyses of Scales Used in the Study

Scale name	Cronbach Alpha	Kaiser-Meyer-Olkin (KMO)	Bartlett
Evaluation Scale for Expert Teacher Career Stage Education	0.986	0.964	4754.361
Evaluation Scale for Expert Teaching Career Stage Exam	0.986	0.960	4906.663
Evaluation Scale for Expert Teaching Career Stage Title	0.996	0.967	6136.009

p<0.001, N=250

Data Collection And Analysis

In order to collect data in the quantitative phase, the addresses and phone numbers of 227 schools in the Selçuklu district of Konya were obtained from the website of the Ministry of National Education (MEB, 2024). Since 37 of these schools were kindergartens with common principals, they were not contacted again. Of the remaining 190 schools, 122 were contacted by phone during working hours and the research details were shared with the school administrators. With the permission of the administrator, the necessary information and surveys to be forwarded to expert teachers were sent via WhatsApp and/or e-mail. In addition, 26 schools were visited in person and surveys were administered with expert teachers during breaks or free class hours. 42 schools could not be reached due to distance or other communication problems.

In the study, scales were applied to a total of 550 expert teachers and SPSS-25.0 package program was used to analyze the obtained data. Before analyzing the data, it was checked whether the scales showed normal distribution. Normal distribution is a type of distribution that is statistically characterized by a certain mean and standard deviation value (Büyükoztürk et al., 2020, p. 57). As a result of the normality assumptions obtained, One-Way ANOVA was used in the analyzes in cases showing normal distribution, and Mann Whitney U, Kruskal Wallis tests and Spearman Correlation Analysis were used in cases not showing normal distribution.

Qualitative Stage

Working Group

The participants of the qualitative phase of the research were determined using the maximum variation sampling method, one of the purposeful sampling approaches. The maximum variation sampling method is used to reveal the main themes covering various differences associated with the event or phenomenon being examined (Erdem and Tutar, 2020). In order to ensure diversity among the expert teachers who were entitled to become expert teachers by taking the career stage exam in the state schools affiliated with the Ministry of National Education in Selçuklu district of Konya province, 30 volunteer expert teachers from different genders, branches, seniority and school levels were selected as participants.

Data Collection Tools

In the qualitative stage of the study, a semi-structured interview form containing five questions was utilized to explore expert teachers' views on the Teaching Career Ladder Application within the context of the General Competencies for the Teaching Profession. The questions were developed with input from a faculty member specializing in curriculum development and finalized through necessary revisions and additions. This interview form was designed to gain an in-depth understanding of expert teachers' perspectives on the career ladder system. It included open-ended questions aimed at assessing teachers' perceptions of the application, its impact on their professional competencies, and the perceived benefits and drawbacks of the process. The items were structured in alignment with the General Competencies for the Teaching Profession and were intended to examine teachers' knowledge, skills, attitudes, and values; their experiences throughout the career ladder process; and the perceived effects of the application on the teaching profession.

The interview form was developed in alignment with the research questions to collect data for the qualitative dimension of the study. The questions were designed to gain an in-depth understanding of teachers' perspectives on the implementation of the expert teacher career step. During the question development process, a literature review was conducted, and interview form samples from previous qualitative studies on the expert teacher implementation, as well as the framework of the General Competencies for the Teaching Profession, were utilized.

To ensure the validity and reliability of the interview form, the questions were shared with a faculty member specialized in curriculum development and two experienced expert teachers. Based on their feedback and suggestions, certain questions were revised for linguistic clarity and simplicity. In addition, a pilot study was conducted with three expert teachers to test the applicability of the interview form. The final version of the form was created based on the feedback obtained from the pilot application.

In the demographic section of the interview form, which was developed with expert input, seven questions were included to collect information such as the teacher's name, gender, age, level of education, years of professional experience, appointment status to the expert teacher career step, and the school/institution they work at.

Specific attention was given to the usefulness of the training materials provided (e.g., workbooks, video lectures), the application's role in supporting professional growth, and its overall strengths and weaknesses as perceived by the participants. For example, the questions used in the interview form; "Did you find the training materials (workbook, video lessons, etc.) presented within the scope of the Expert Teacher Training Program sufficient and useful?", "Do you think the Teaching Career Ladder application contributed to your professional knowledge, skills, and competencies?", "What are the advantages and disadvantages of this application for the teaching profession?" can be given.

Data Collection and Analysis

Data for the study were gathered through in-person interviews with 30 expert teachers employed at public schools in the Selçuklu district of Konya, all of whom had successfully passed the expert teacher examination and earned the corresponding title. Participation was entirely voluntary, and all individuals signed an informed consent form approved by the Ministry of National

Education prior to the interviews. With participants' permission, the interviews were audio-recorded in quiet settings to ensure clarity and accuracy.

For qualitative data analysis, Nvivo12 Plus program was used to examine the experiences and opinions of expert teachers regarding career ladder applications. NVivo is a program that allows organizing data under codes, establishing relationships between these codes, visualizing and reporting data (Cassell et al., 2005). In order to ensure confidentiality in the research, expert teachers were coded as UÖ1, UÖ2, ... UÖ30. The coding process was carried out using the NVivo 12 software. After transcribing participants' responses verbatim, the written texts were uploaded to the program and coded according to their content. NVivo's tools were utilized for visualizing the codes, conducting frequency analysis, and identifying relationships between themes.

Validity refers to how accurately a scale or research tool measures the feature it aims to measure, independent of other features (Büyüköztürk et al., 2020, p. 121).

Information on the reliability of the research is discussed under the heading of data collection tools. In this section, the issue of validity is particularly emphasized. In the quantitative research, a representative group was selected from a large universe using cluster and criterion sampling methods, thus increasing the level of validity. In the qualitative research, in-depth interviews were conducted with 30 teachers with different experiences and perspectives using the maximum variation sampling method. In order to increase validity in quantitative and qualitative research, various techniques such as data, researcher, method and sample diversity are used (Tutar, 2022). In terms of external validity, while the large and representative sample used in quantitative research increases the generalizability of the findings, the maximum variation sampling method used in qualitative research may limit generalizability. Therefore, it is important to evaluate the findings as valid only in the context in which the research was conducted (Can, 2022).

Findings and Interpretation

In this section, the quantitative and qualitative findings regarding the sub-problems of the research are given. The findings are presented in a sequence in accordance with the determined objectives of the research.

Analysis and Interpretations of Quantitative Data

Quantitative Findings on Teacher Opinions Regarding the Evaluation of Expert Teacher Career Stage Education

In the research, the first sub-problem determined as "To what extent do expert teachers think that the training program taken before the teaching career stage exam contributes to the General Competencies of the Teaching Profession?" was answered. In this section, first of all, descriptive statistics regarding the total scale score and item-based scores received by expert teachers from the Expert Teacher Career Stage Education Evaluation Scale were given. Then, the results regarding the normality assumptions made in order to determine whether the groups were normally distributed or not, and whether this problem differed significantly according to gender and branch variables, whether there was a significant relationship between exam scores and also seniority were determined, and the findings obtained from the quantitative data obtained to answer the problems were reached.

The total scores of the expert teachers participating in the study on the Evaluation Scale for Expert Teacher Career Step Education were calculated as $\bar{X}=25.96$ for the evaluation of the Expert Teacher Career Step Education. When the scores obtained from the scale were examined, it was seen that the minimum and maximum values were between 1 and 5, the standard deviation varied between 1.13 and 1.17, and the average scores were between 2.30 and 2.41. In this direction, it was determined that the participating expert teachers gave a score of "very little contribution" in the evaluation of the Expert Teacher Career Step Education. The teachers stated that the training and video lessons contributed to the knowledge of the legislation at a moderate level but made a very limited contribution to the subjects of field knowledge, teaching planning, creating a learning environment, managing teaching processes, measurement and evaluation, values education, and self-evaluation. In addition, they think that they did not contribute to field education and effective communication skills.

Normality Assumptions and Difference Analysis Findings Regarding the Gender Variable of The Expert Teacher Career Stage Education Evaluation Scale

The normality test results for the gender variable of the Expert Teacher Career Stage Education Evaluation Scale scores used in the research are given in Table 2.

Table 2. Results of Normality Test of Expert Teacher Career Stage Education Evaluation Scale Scores for Gender Variable

Gender	Kolmogorov-Smirnov			Shapiro-Wilk's		
	Statistics	df	p	Statistics	df	p
Woman	.122	264	.000	.949	264	.000
Male	.099	280	.000	.938	280	.000

p<.05

As seen in Table 2, as a result of the analysis, it was seen that the total scores of the Expert Teacher Career Stage Education Evaluation Scale on the gender variable did not show a normal distribution according to the p<.05 value. In order to determine whether the opinions of expert teachers regarding the contribution of the training program taken before the teaching career stage exam to the General Competencies of the Teaching Profession differ significantly according to gender, the Mann-Whitney U Test,

which is a non-parametric test, was used since there were two independent categories and the analysis results are given in Table 3.

Table 3. Mann-Whitney U Test Results on Whether the Evaluation Scores of Expert Teacher Career Stage Education Differ According to the Gender Variable

Gender	N	Average Rank	Total Rank	U	P
Woman	268	254.25	68138	32092	.002*
Male	282	295.40	83387		
Total	550				

*p<.05

According to Table 3, there is a significant difference between male and female participants in terms of the contribution of the training program received before the teaching career stage exam to the general competencies of the teaching profession (U=32092, p<.05). Male participants received higher scores than female participants in terms of the contribution of this training.

Normality Assumptions and Difference Analysis Findings of the Branch Variable of The Expert Teacher Career Stage Education Evaluation Scale

Evaluation Scale for Expert Teacher Career Stage Education the normality test results of the branch variable of the scores were examined and the results are given in Table 4.

Table 1. Normality Test Results of Expert Teacher Career Stage Education Evaluation Scale Scores for Branch Variable

Branch	Kolmogorov-Smirnov			Shapiro-Wilk's		
	Statistics	df	p	Statistics	df	p
Basic Education Group	.145	219	.000	.930	219	.000
Numerical Group	.101	114	.007	.947	114	.000
Social Group	.092	71	.200	.946	71	.004
Talent Group	.126	37	.148	.917	37	.009
Occupation Group	.121	17	.200	.943	17	.360
Language Group	.113	89	.007	.922	89	.000

p<.05

When Table 4 is examined, it was determined that the subgroups of the scale applied to measure the contribution of the training provided to expert teachers before the career stage exam to the general competencies of the teaching profession showed a normal distribution and in the analyses of the Expert Teacher Career Stage Training Evaluation Scale to be made at the branch level, one-way analysis of variance (One-Way ANOVA) test was used and the analysis results are given in Table 5.

Table 2. ANOVA Results on Whether the Evaluation of Expert Teacher Career Stage Education Shows Differentiation According to the Branch Variable

Source of Variance	KT	df	KO	f	p.
Intergroup	984,682	5	196,936	1,473	.197
Within groups	72711.356	544	133,661		
Total	73696.038	549			

p<.05

When Table 5 is examined, it is seen that the opinions regarding the contribution of the pre-exam training to the general competencies of the teaching profession are close among all branches as a result of the analysis and no significant difference was found.

Normality Assumptions and Correlation Analysis Findings for the Exam Score Variable of The Expert Teacher Career Stage Education Evaluation Scale

The normality test results for the exam score variable of the Expert Teacher Career Stage Education Evaluation Scale scores used in the research are given in Table 6.

Table 3. Normality Test Results for Evaluating Expert Teacher Career Stage Education Scale Scores and Exam Score Variable

Variable	Kolmogorov-Smirnov			Shapiro-Wilk's		
	Statistics	df	p	Statistics	df	p
Scale Score	.114	550	.000	.937	550	.000
Exam Score	.224	550	.000	.794	550	.000

p<.05

When Table 6 is examined, it is seen that the data regarding the scale score and exam score are normal. It was observed that the data were not distributed and Spearman's Rank Correlation Coefficient analysis was used in the analyses related to the exam scores variable of the Expert Teacher Career Stage Education Evaluation Scale and the results of the analyses are given in Table 7.

Table 4. Descriptive Statistics and Spearman Correlations on Whether There is a Significant Relationship Between Evaluation Scores and Exam Scores for Expert Teacher Career Stage Education

Variable	N	Avg.	SS	Scale Score	Exam Score	p
Scale Score	550	25.96	11.58	-	-.027	.530
Exam Score	550	95.35	4.61	-.027	-	.530

$p < .05$

When Table 7 is examined, according to Spearman's Correlation Coefficient analysis; a negative and very weak relationship was found between the opinions of the participating expert teachers and the exam scores depending on the education received before the career stage exam ($r_{\text{spearman}} = -.027$, $p = .530$). Since this value is very close to 0, it can be found that the education received does not directly affect the exam success within the framework of the general competencies of the teaching profession and there is no strong connection between these variables.

Normality Assumptions and Relationship Analysis Findings of the Seniority Variable of The Expert Teacher Career Stage Education Evaluation Scale

The normality test results of the seniority variable of the Expert Teacher Career Stage Education Evaluation Scale scores used in the study are given in Table 8.

Table 5. Results of Normality Test of Seniority Variable of Scale Scores for Evaluating Expert Teacher Career Stage Education

Variable	Kolmogorov-Smirnov			Shapiro-Wilk's		
	Statistics	df	p	Statistics	df	p
Scale Score	.114	550	.000	.937	550	.000
Seniority	.074	550	.000	.978	550	.000

$p < .05$

When Table 8 is examined, it is seen that the data regarding the scale score and seniority are not normally distributed. In this direction, Spearman's Rank Correlation Coefficient analysis was used in the analysis regarding whether there is a significant relationship between the contribution levels (scores) of the training program taken before the teaching career stage exam to the general competencies of the teaching profession depending on the seniority of the teachers participating in the research and the results of the analyses are summarized in Table 9.

Table 6. Descriptive Statistics and Spearman Correlations on Whether There is a Significant Relationship Between the Evaluation Scores of Expert Teacher Career Stage Education and Their Seniority

Variable	N	Avg.	SS	Scale Score	Seniority	p
Scale Score	550	25.96	11.58	-	.070	.103
Seniority	550	21.17	6.42	.070	-	.103

$p < .05$

When Table 9 is examined, according to Spearman's Correlation Coefficient analysis; a positive but very weak relationship was found between the opinions of the participating expert teachers and their seniority depending on the training received before the career stage exam ($r_{\text{spearman}} = .070$, $p = .103$). In addition, this relationship is not considered statistically significant since the p value is greater than 0.05. This situation reveals that there is no significant relationship between the seniority of the expert teachers and the training evaluation scores they receive, and that seniority is not a determining factor in the participation of expert teachers in the training process and the benefit they obtain from this process.

Quantitative Findings on Teachers' Views on Evaluating The Expert Teaching Career Stage Exam

In this section of the research, the answer to the second sub-problem determined as "To what extent do expert teachers think that the Teaching Career Stage Exam can measure the General Competencies of the Teaching Profession?" was sought. In this section, first of all, descriptive statistics regarding the total score and item-based scores of the Expert Teachers from the Expert Teacher Career Stage Exam Evaluation Scale were given. Then, the results regarding the normality assumptions made in order to determine whether the groups were normally distributed and whether this problem differed significantly according to gender and branch variables, whether there was a significant relationship between the exam scores and also their seniority were determined and the findings obtained from the quantitative data obtained to answer the problems were reached.

The total score of the expert teachers participating in the study, obtained from the Expert Teacher Career Stage Exam Evaluation Scale, was calculated as =23.71 for the evaluation of the Expert Teacher Career Stage Exam. When the score ranges of the scale for the Expert Teacher Career Stage Exam evaluation are examined, it is seen that the minimum and maximum values are between 1 and 5, the standard deviation varies between 1.07 and 1.12, and the average scores are between 1.98 and 2.24. In this direction, it was determined that the participating expert teachers gave a score at the level of "definitely does not measure"

in the evaluation of the Expert Teacher Career Stage Exam. $\bar{X}$ Teachers stated that the exam did not adequately evaluate field knowledge, field education, knowledge of legislation and pedagogical competencies. They also stated that national, moral and universal values, effective communication and cooperation skills and competencies that support personal and professional development were definitely not evaluated by the exam. It is thought that only the skills of creating a learning environment and measurement-evaluation are measured very little.

Normality Assumptions and Difference Analysis Findings Regarding the Gender Variable of The Expert Teacher Career Stage Exam Evaluation Scale

The normality test results for the gender variable of the Expert Teaching Career Stage Exam Evaluation Scale scores, which is the second scale used in the study, are given in Table 10.

Table 7. Results of Normality Test of Expert Teaching Career Stage Exam Evaluation Scale Scores for Gender Variable

Gender	Kolmogorov-Smirnov			Shapiro-Wilk's		
	Statistics	df	p	Statistics	df	p
Woman	.139	264	.000	.912	264	.000
Male	.124	280	.000	.916	280	.000

$p < .05$

As seen in Table 10, as a result of the analyses, it was determined that the total scores of the Expert Teacher Career Stage Exam Evaluation Scale did not show a normal distribution on the gender variable, since $p < .05$. Whether there was a significant difference in the levels (scores) of the teaching career stage exam measuring general competencies of the teaching profession depending on the gender of the teachers participating in the research was examined with the Mann-Whitney U Test, one of the non-parametric tests, and the results are summarized in Table 11.

Table 8. Mann-Whitney U Test Results on Whether the Evaluation of the Expert Teaching Career Stage Exam Differs According to the Gender Variable

Gender	N	Average Rank	Total Rank	U	P
Woman	268	260.37	69779	33733	.028*
Male	282	289.88	81746		
Total	550				

* $p < .05$

According to the Mann-Whitney U test results, since the p value is .028 ($p < .05$), there is a statistically significant difference in the opinions of male and female teachers regarding the level of measuring the general competencies of the teaching profession in the career ladder exam. The mean rank of male teachers (289.88) is higher than that of female teachers (260.37), which shows that male teachers evaluate the exam more positively and that gender is a factor in the perception of the exam.

Normality Assumptions and Difference Analysis Findings of The Branch Variable of The Expert Teacher Career Stage Exam Evaluation Scale

Evaluation Scale for Expert Teaching Career Stage Exam the normality test results of the branch variable of the scores were examined and the results are given in Table 12.

Table 9. Results of Normality Test of Expert Teaching Career Stage Exam Evaluation Scale Scores for Branch Variable

Branch	Kolmogorov-Smirnov			Shapiro-Wilk's		
	Statistics	df	p	Statistics	df	p
Basic Education Group	.153	219	.000	.911	219	.000
Numerical Group	.158	114	.000	.889	114	.000
Social Group	.120	71	.013	.919	71	.000
Talent Group	.161	37	.017	.898	37	.003
Occupation Group	.192	17	.095	.880	17	.032
Language Group	.120	89	.003	.919	89	.000

$p < .05$

When Table 12 is examined, Shapiro-Wilk's test was evaluated for groups smaller than 50, and Kolmogorov-Smirnov test was evaluated for groups larger than 50, and it was seen that the groups did not show a normal distribution according to the results obtained ($p < .05$). Since there were more than three groups in the branch variable analysis of the Expert Teacher Career Stage

Exam Evaluation Scale, assuming that they were not normally distributed, the Kruskal Wallis-H test, one of the non-parametric tests, was used and the results are given in Table 13.

Table 10. Kruskal Wallis-H Test Results on Whether the Evaluation of the Expert Teaching Career Stage Exam Differs According to the Branch Variable

Branch	N	Average Rank	df	X ²	p
Basic Education Group	219	285.93	5	11	.102
Numerical Group	122	242.05			
Social Group	63	289.18			
Talent Group	40	290.76			
Occupation Group	17	230.71			
Language Group	89	287.70			

p<.05

According to Table 13, according to the Kruskal-Wallis H test results, there is no statistically significant difference between the evaluation scores of the Expert Teacher Career Stage Exam and the branch groups (p>.05). This shows that the exam evaluations of the teachers are similar according to their branches and the branch variable is not a determining factor.

Normality Assumptions and Correlation Analysis Findings Regarding the Exam Score Variable of The Expert Teacher Career Stage Exam Evaluation Scale

In this section, normality tests were performed for the exam score variable of the Expert Teacher Career Stage Exam Evaluation Scale scores used in the research and the results are given in Table 14.

Table 11. Results of the Normality Test of the Expert Teaching Career Stage Exam Evaluation Scale Scores for the Exam Score Variable

Variable	Kolmogorov-Smirnov			Shapiro-Wilk's		
	Statistics	df	p	Statistics	df	p
Scale Score	.134	550	.000	.914	550	.000
Exam Score	.224	550	.000	.794	550	.000

p<.05

When Table 14 is examined, it is determined that the data regarding the scale score and exam score do not show a normal distribution. For this reason, Spearman's Rank Correlation Coefficient analysis was preferred in the analyses regarding the exam scores variable of the Expert Teaching Career Stage Exam Evaluation Scale. The results of this analysis are presented in Table 15.

Table12. Descriptive Statistics and Spearman Correlations on Whether There is a Significant Relationship Between Expert Teaching Career Stage Exam Evaluation Scores and Exam Scores

Variable	N	Avg.	SS	Scale Score	Exam Score	p
Scale Score	550	23.71	11.14	-	-.089	.038*
Exam Score	550	95.35	4.61	-.089	-	.038*

*p<.05

According to the data presented in Table 15, there is a negative correlation between scale scores and exam scores ($r_{\text{spearman}} = -.089$, $p=.038$). This shows that there is a weak but statistically significant relationship between scale scores and exam scores. However, the low correlation coefficient shows that this relationship is weak and its reflection on practice and therefore professional competencies may be limited.

Normality Assumptions and Relationship Analysis Findings of The Seniority Variable of The Expert Teacher Career Stage Exam Evaluation Scale

The results of the normality tests performed on the Expert Teacher Career Stage Exam Evaluation Scale scores used in the study according to the seniority variable are given in Table 16. The results obtained show whether the distribution of scale scores at the seniority level is in accordance with the normality assumptions.

Table 13. Results of Normality Test of Seniority Variable of Evaluation Scale Scores of Expert Teaching Career Stage Exam

Variable	Kolmogorov-Smirnov			Shapiro-Wilk's		
	Statistics	df	p	Statistics	df	p
Scale Score	.134	550	.000	.914	550	.000
Seniority	.074	550	.000	.978	550	.000

$p < .05$

When Table 16 is examined, it is seen that the data regarding scale score and seniority do not show a normal distribution. Therefore, Spearman Rank Correlation Coefficient analysis was used to examine whether there is a significant relationship between the contribution levels of the teaching career stage exam to the general competencies of the teaching profession according to the seniority of the teachers participating in the research. The results of these analyses are summarized in Table 17.

Table 14. Descriptive Statistics and Spearman Correlations on Whether There is a Significant Relationship Between Expert Teacher Career Stage Exam Evaluation Scores and Seniority

Variable	N	Avg.	SS	Scale Score	Seniority	p
Scale Score	550	23.71	11.14	-	.075	.078
Seniority	550	21.17	6.42	.075	-	.078

$p < .05$

According to the data presented in Table 17, there is a very weak positive correlation between the scale score and seniority ($r_{\text{spearman}} = .075$, $p = .078$). However, this relationship is not statistically significant ($p > .05$). This shows that although the scale scores tend to increase as the seniority of the teachers increases, this increase is not at a significant level. In other words, there is no strong relationship between the seniority of the teachers and the scale scores they receive.

Quantitative Findings on Teachers' Opinions Regarding the Evaluation of The Title of Expert Teacher

In the research, an answer was sought to the third problem determined as "To what extent do expert teachers think that receiving the title of expert teacher contributes to the General Competencies of the Teaching Profession?" First of all, descriptive statistics were given regarding the total score and item-based scores of expert teachers from the Expert Teacher Title Evaluation Scale. Then, the results regarding the normality assumptions made in order to determine whether the groups were normally distributed or not, and whether this problem differed significantly according to gender and branch variables, whether there was a significant relationship between exam scores and also seniority were determined, and the findings obtained from the quantitative data obtained to answer the problems were reached.

The total scores of the expert teachers participating in the study on the Expert Teacher Title Evaluation Scale were calculated as $\bar{X} = 22.4$ for the evaluation of the career step title. When the score ranges of the scale for the evaluation of the career step title were examined, it was determined that the minimum and maximum values were between 1 and 5, the standard deviation varied between 1.07 and 1.13, and the average scores were between 1.86 and 2.11. Accordingly, it was observed that the participating expert teachers gave a score of "did not contribute at all" to all items regarding the evaluation of the expert teacher title. The findings show that the expert teacher title did not provide the expected contribution in developing the professional competencies of teachers and that teachers did not benefit sufficiently from this title.

Normality Assumptions and Difference Analysis Findings Regarding the Gender Variable of The Expert Teacher Title Evaluation Scale

The normality test results for the gender variable of the Expert Teacher Title Evaluation Scale scores used thirdly in the study are given in Table 18.

Table 15. Normality Test Results of Expert Teacher Title Evaluation Scale Scores for Gender Variable

Gender	Kolmogorov-Smirnov			Shapiro-Wilk's		
	Statistics	df	p	Statistics	df	p
Woman	.163	264	.000	.869	264	.000
Male	.162	280	.000	.881	280	.000

$p < .05$

When Table 18 is examined, it is concluded that the groups do not show a normal distribution because the total scores of the Expert Teacher Title Evaluation Scale are $p < .05$. For this reason, the Mann-Whitney U Test, which is one of the non-parametric tests, was used in the analyses of the Expert Teacher Title Evaluation Scale regarding the gender variable, and the results of this analysis are given in Table 19.

Table 16. Mann-Whitney U Test Results on Whether the Evaluation of the Expert Teacher Title Differs According to the Gender Variable

Gender	N	Average Rank	Total Rank	U	P
Woman	268	261.33	70036	33990	.039*
Male	282	288.97	81489		
Total	550				

*p<.05

When Table 19 is examined, there is a significant difference between male and female expert teachers regarding the contribution of receiving the title of expert teacher to the general competencies of the teaching profession according to gender (U=33990, p<.05). This difference is in the direction of male participants. It is seen that the mean rank of male participants regarding the contribution of receiving the title of expert teacher to the general competencies of the teaching profession is higher than that of female participants.

Normality Assumptions and Difference Analysis Findings of The Branch Variable of The Expert Teacher Title Evaluation Scale

Normality test results of the Expert Teaching Title Evaluation Scale scores, which is the third scale used in the study, for the branch variable are given in Table 20.

Table 17. Results of Normality Test of Expert Teacher Title Evaluation Scale Scores on Branch Variable

Branch	Kolmogorov-Smirnov			Shapiro-Wilk's		
	Statistics	df	p	Statistics	df	p
Basic Education Group	.175	219	.000	.873	219	.000
Numerical Group	.180	114	.000	.850	114	.000
Social Group	.160	71	.000	.874	71	.000
Talent Group	.165	37	.012	.875	37	.001
Occupation Group	.183	17	.135	.902	17	.072
Language Group	.176	89	.000	.855	89	.000

p<.05

When Table 20 is examined, Shapiro-Wilk's test was evaluated for groups smaller than 50, and Kolmogorov-Smirnov test was evaluated for groups larger than 50, and it was seen that the groups did not show a normal distribution according to the results obtained (p<.05). Based on this result, Kruskal Wallis-H test, one of the non-parametric tests, was used in the analyses of the Expert Teacher Title Evaluation Scale's branch variable, and the results of the test are given in Table 21.

Table 18. Results of Kruskal Wallis-H Test on Whether the Evaluation of Expert Teacher Title Differs According to the Branch Variable

Branch	N	Average Rank	df	X ²	p
Basic Education Group	219	284.29	5	12.84	.225
Numerical Group	122	248.05			
Social Group	63	292.57			
Talent Group	40	305.19			
Occupation Group	17	275.47			
Language Group	89	266.10			

p<.05

According to the data in Table 21, the results of the Kruskal Wallis-H test conducted between the branch groups show that there is no statistically significant difference between the branches in the evaluation of the Expert teacher title (p>.05). The results obtained reveal that the evaluations of the teachers regarding the Expert teacher title do not differ according to their branches.

Normality Assumptions and Correlation Analysis Findings Regarding the Exam Score Variable of The Expert Teacher Title Evaluation Scale

Normality tests were performed for the exam score variable of the Expert Teacher Title Evaluation Scale scores used in the research and the results are given in Table 22.

Table 19. Normality Test Results of the Expert Teacher Title Evaluation Scale Scores for the Exam Score Variable

Variable	Kolmogorov-Smirnov			Shapiro-Wilk's		
	Statistics	df	p	Statistics	df	p
Scale Score	.161	550	.000	.871	550	.000
Exam Score	.224	550	.000	.794	550	.000

$p < .05$

When Table 22 is examined, it is determined that the data regarding the scale score and exam scores do not show a normal distribution. Considering this situation, Spearman Rank Correlation Coefficient analysis was used in the analyses regarding the exam scores variable of the Expert Teaching Career Stage Exam Evaluation Scale and the results of the analysis are presented in Table 23.

Table 20. Descriptive Statistics and Spearman Correlations on Whether There is a Significant Relationship Between Expert Teacher Title Evaluation Scores and Seniority

Variable	N	Avg.	SS	Scale Score	Exam Score	p
Scale Score	550	22.4	11.51	-	.059	.169
Exam Score	550	95.35	4.61	.059	-	.169

$p < .05$

According to the data presented in Table 23, it was examined whether there was a significant relationship between the Expert Teacher Title Evaluation Scores and the exam scores. The results show that there is a very weak positive correlation between the scale score and the exam score ($r_{\text{spearman}} = .059$, $p = .169$). However, this correlation coefficient is not statistically significant ($p > 0.05$). This indicates that there is no significant relationship between the scale score and the exam score.

Normality Assumptions and Correlation Analysis Findings of The Seniority Variable of The Expert Teacher Title Evaluation Scale

The results of the normality tests performed on the Expert Teacher Title Evaluation Scale scores used in the study according to the seniority variable are given in Table 24.

Table 21. Results of Normality Test of Seniority Variable for Evaluation Scale Scores of Expert Teacher Title

Variable	Kolmogorov-Smirnov			Shapiro-Wilk's		
	Statistics	df	p	Statistics	df	p
Scale Score	.161	550	.000	.871	550	.000
Seniority	.074	550	.000	.978	550	.000

$p < .05$

When looking at the table, it is seen that the data regarding the scale score and seniority do not show a normal distribution, therefore, Spearman Rank Correlation Coefficient analysis was used to understand whether there is a significant relationship between the contribution levels of the teaching title to the general competencies of the teaching profession according to the seniority of the teachers participating in the research. The analysis results are summarized in Table 25.

Table 22. Descriptive Statistics and Spearman Correlations on Whether There is a Significant Relationship Between Expert Teacher Title Evaluation Scores and Seniority

Variable	N	Avg.	SS	Scale Score	Seniority	p
Scale Score	550	22.4	11.51	-	.059	.169
Seniority	550	21.17	6.42	.059	-	.169

$p < .05$

According to the data given in Table 25, there is a very weak positive correlation between the scale score and seniority ($r_{\text{spearman}} = .059$, $p = .169$). However, this relationship was not found to be statistically significant ($p > .05$). This shows that as the seniority of the teachers increases, the scale scores increase, but this increase is not at a significant level. In other words, there is no significant relationship between the seniority of the teachers and the scale scores.

Analysis and Interpretations of Qualitative Data

Findings Obtained from Qualitative Data Regarding General Thoughts About Teaching Career Ladder Application

In order to find an answer to the question "What are the experiences and opinions of expert teachers regarding the teaching career ladder application?" in the qualitative research phase, the general opinions of expert teachers participating in the research regarding the teaching career ladder application were divided into 4 main themes. A total of 14 different sub-themes were reached

from these 4 main themes. The frequency and participant distributions of the themes and sub-themes reached regarding these opinions are given in Table 26.

Table 23. Findings Regarding General Thoughts on the Application of Teaching Career Step

Theme	Child Theme	f	Participants
Positive Thoughts	Personal Development and professional skill increase	3	UÖ4, UÖ6, UÖ14
	Innovative learning opportunities	2	UÖ6, UÖ16
	Career development and planning	3	UÖ7, UÖ11, UÖ25
Negative Thoughts	The futility of the exam	10	UÖ3, UÖ8, UÖ13, UÖ15, UÖ21, UÖ22, UÖ23, UÖ25, UÖ27, UÖ29
	Content and structure of the exam	5	UÖ9, UÖ10, UÖ16, UÖ17, UÖ27
	Financial concerns	2	UÖ2, UÖ5
	Incompatibility with the nature of the teaching profession	6	UÖ1, UÖ18, UÖ20, UÖ26, UÖ28, UÖ30
Advantages of the teaching career stage	Pecuniary advantage	13	UÖ1, UÖ2, UÖ5, UÖ7, UÖ11, UÖ13, UÖ17, UÖ19, UÖ23, UÖ25, UÖ26, UÖ27, UÖ29
	Professional Development, Knowledge and Skill Acquisition	12	UÖ1, UÖ6, UÖ8, UÖ9, UÖ10, UÖ14, UÖ15, UÖ16, UÖ18, UÖ22, UÖ29, UÖ30
	Increased Motivation and Interest	5	UÖ6, UÖ8, UÖ15, UÖ16, UÖ30
	Professional Prestige	2	UÖ4, UÖ20
Disadvantages of the teaching career stage	Inequality and Discrimination	9	UÖ1, UÖ2, UÖ6, UÖ8, UÖ12, UÖ17, UÖ21, UÖ27, UÖ29
	Stress and Loss of Self-Esteem	2	UÖ12, UÖ21
	Concerns about Material Inequality and Loss of Rights	7	UÖ1, UÖ2, UÖ3, UÖ6, UÖ17, UÖ27, UÖ29
	Waste of Time	2	UÖ23, UÖ24
	Status Differences	7	UÖ2, UÖ7, UÖ8, UÖ12, UÖ17, UÖ21, UÖ27

As seen in Table 26, opinions on the teaching career ladder application indicate more than one factor. In Table 26, within the scope of the findings regarding the General Thoughts on the Teaching Career Ladder Application, the sub-themes, frequencies and participant codes regarding the main theme of “*Positive Thoughts*” are seen. When the codes are examined, it is seen that this main theme is divided into three sub-themes as “*Personal Development and Increase in Professional Skills*” (f: 3), “*Innovative Learning Opportunities*” (f: 2) and “*Career Development and Planning*” (f: 3).

Personal development and increased professional skills

The view that teachers need to develop themselves professionally and that this practice increases personal and professional skills is gathered under this sub-theme. Some of the teachers thought that there should be more effective, long-term trainings to be motivated and to complete the deficiencies, and that this training should be supported with practical applications, and they found the practice useful and stated that the practice allowed them to complete the deficiencies in certain areas [10%].

UÖ4: *"I think there could have been more effective training instead of these, that's my opinion. When we say more effective training, for example, we took a first aid seminar, none of us objected because we needed it in our daily lives."*

UÖ14: *"I think it would be healthier if it were spread over a longer-term program, and I personally think it would be more appropriate if it were supported by practical applications rather than remaining theoretical."*

Innovative learning opportunities

It has been suggested that the exam system of the application motivates teachers, encourages continuous learning and that the educational content is of high quality. It has also been stated that exams are a source of motivation for teachers, but teachers who rely on their experience can neglect learning over time, therefore exams play an important role in keeping the continuous learning process alive [6.67%].

UÖ6: *"Well, I realized that exams are a very good thing to motivate teachers. You know, normally... We left it too long, right? Yes. I can't say that I'm not open to learning, but... After that, you really stop learning by relying on experience."*

Career development and planning

It was stated that it is important for teachers to have a certain career stage in their professional lives, but that it is not right to evaluate this process only with an exam. It was emphasized that the application of career stages is necessary, but the appointment of teachers to these stages should be based on seniority, that is, the years spent in the profession, rather than an exam. It was also stated that experience and working hours should be taken as basis in appointments [10%].

UÖ11: *"So, being an expert teacher is already based on experience. You become an expert teacher, that is, as the years go by, you develop in terms of your command, classroom management, teacher-student relationship, parent relationship, etc. That's why career steps can be fixed to the year, not to the exam."*

UÖ25: *"I think there should be a career ladder. But it seems more appropriate to do this in relation to the duration of teachers' employment rather than the exam."*

"Negative Thoughts" When the sub-themes, frequencies and participant codes related to the main theme are examined, it is seen that this main theme is divided into four sub-themes: *"Unnecessity of the exam"* (f: 10), *"Content and structure of the exam"* (f: 5), *"Financial concerns"* (f: 2) and *"Incompatibility with the nature of the teaching profession"* (f: 6).

In the research, when the participants were asked about their thoughts on the teaching career ladder application, some teachers responded negatively to the question that there were some injustices during the exam process, there were stressful aspects, the exam content was inadequate and inappropriate, while some teachers argued that professional knowledge and skills cannot be measured by exams and that this form of evaluation was not an approach suitable for the nature of teaching.

The futility of the exam

Some teachers argue that the teaching profession is inherently specialized and therefore there is no need for an additional exam. They find it unnecessary to determine career ladders based on exams and state that this practice creates a hierarchical distinction among teachers. However, some teachers stated that the transition to career ladders should be carried out systematically and fairly, that exams are unnecessary and create additional stress and burden on teachers [33.3%].

UÖ13: *"I don't think it's appropriate to evaluate these with an exam. I mean, as a teacher who has worked for 20 years, I think I have sufficient command of these. I passed the exam successfully without much difficulty in the questions that came up. I mean, the training I took, the courses I took, the exam I took afterwards didn't affect me that much, they didn't affect me extra. There weren't any subjects that helped me develop, let me say that."*

UÖ15: *"I don't think it's a necessary exam. Frankly, teachers are a bit stressed and burdened. It was a stressful process for us, since it was our first time taking the exam, there were teachers who were stressed. I was a bit more relaxed, though. But I think there was a burden on the other side."*

Content and structure of the exam

Criticisms have been voiced that the exam has various deficiencies in terms of measurement and evaluation and that it is not directly related to education. Teachers state that the exam questions do not truly measure professional competence and that some topics are unnecessary. In addition, it has been suggested that this exam is not suitable for teachers with a certain level of experience, is inadequate in terms of measuring professional skills and does not provide development in field knowledge [16.6%].

UÖ9: *"It has nothing to do with my field. They provided general training in terms of teaching, but it did not include anything related to my field."*

UÖ27: *"I think this study was rushed. Its content should have been much more comprehensive. It should have been handled on a branch basis."*

Financial concerns

Some teachers stated that the main purpose of the exam for career steps is to gain financial gain and that the majority of teachers are involved in this process with economic expectations. In addition, it was stated that teachers participate in the exam in order to improve their living standards due to financial concerns and that this practice is carried out to provide economic support to teachers. In this context, evaluations were also encountered that the exam is used as a tool [6.6%].

UÖ2: *"Most of the teachers who took the exam actually got involved in this process because of financial concerns. Because we all do the same job, no one does anything different. But there were those who wanted to make a difference economically by taking the exam."*

Incompatibility with the nature of the teaching profession

Some of the participants stated that the teaching profession is essentially a process based on knowledge and experience and that the examination system is not compatible with the nature of this profession. It was emphasized that competence is gained over time, and therefore the exam is inadequate in measuring professional development. In addition, opinions were expressed that the teaching profession requires expertise by nature, that it is not right to create hierarchical distinctions among teachers, and that the exam is an element that damages the reputation of the profession. In this context, it was stated that teaching is a profession based on competence and experience, and therefore it cannot be measured by exams, and that this practice is against the nature of the teaching profession [20%].

UÖ1: *"I think it is unnecessary and wrong to implement it. Teaching is already a profession that requires expertise. I do not think it is right to categorize teachers with class distinctions such as expert teachers and head teachers."*

UÖ30: *"Now, let me put it this way, teaching is a profession of love. It is not a profession that you can just look at the paper or you can stereotype by saying, I don't know, there has to be this or that competence. First of all, you have to love teaching. Second, I have learned this for myself in my life as a teacher. We will not do to our students what we would not want to happen to our own children. Therefore, it cannot be measured by exams."*

"Advantages of the Teaching Career Stage" When the sub-themes, frequencies and participant codes related to the main theme are examined, it is seen that this main theme is divided into four sub-themes: *"Material Advantage"* (f: 13), *"Professional Development, Acquisition of Knowledge and Skills"* (f: 12), *"Increase in Motivation and Interest"* (f: 5) and *"Professional Prestige"* (f: 2).

Pecuniary advantage

A large portion of teachers state that one of the most important advantages of the career ladder application is financial gain. This situation, which is seen as an economic improvement, is especially prominent for teachers who have been in the profession for many years. However, some teachers state that the application does not provide them with any significant contribution other than this advantage [43%].

UÖ5: *"There are financial advantages. I don't think it's very important for many others, for someone like me who has been in my profession for 25 years and is approaching retirement."*

UÖ25: *"It didn't provide much of an advantage, I mean there wasn't anything like that, I mean it didn't provide any advantage other than the financial advantage."*

Professional development, knowledge and skills acquisition

Some teachers state that the training provided during the career stage supports professional development, contributes to refreshing existing knowledge and following current developments. In this direction, it is stated that the training improves the professional competence of teachers by increasing their knowledge and skills [40%].

UÖ10: *"I mean, it would be a lie if I said it didn't provide any advantage. Visuality is very important. In some subjects, especially in terms of measurement and evaluation in training, social interaction and communication, some videos were very good. I also think it was beneficial."*

Increased motivation and interest

Some teachers stated that the career ladder process motivated them academically and professionally and contributed to their acquisition of new information during the exam process. This situation increased the interest and commitment of teachers to their profession and allowed them to enter a more conscious learning process. There are also participants who think that the career ladder process increases motivation and professional commitment [16.6%].

UÖ15: *"So it motivated us a little bit during that process. During that process, there was little effort in terms of working. We gained new information."*

Professional prestige

Some teachers stated that the titles obtained within the scope of career steps increase the social status and professional reputation of teachers. It is stated that the titles of expert teacher and head teacher in particular provide trust and respect in the eyes of parents and students. In this context, there are also participants who believe that the titles obtained through the application of teaching career steps provide prestige, trust and respect [6.6%].

UÖ20: *"It has more financial benefits but also provides a little more prestige both in society and in front of students and parents, and since we can say I am an expert or a head teacher, it creates a slightly different perspective in the parents, students or society."*

"Disadvantages of the Teaching Career Stage" When the subthemes, frequencies and participant codes related to the main theme are examined, it is seen that this main theme is divided into five subthemes as *"Inequality and Discrimination"* (f: 9), *"Stress and Loss of Self-esteem"* (f: 2), *"Material Inequality and Concern about Loss of Rights"* (f: 7), *"Waste of Time"* (f: 2) and *"Status Differences"* (f: 7). The statements of the teachers reveal that the career ladder application undermines equality in education, causes separation among teachers and may negatively affect professional motivation and self-esteem.

Inequality and discrimination

Some teachers think that the career ladder practice leads to inequality and discrimination among their colleagues. It has been stated that differences in title cause a status-based separation among teachers, that some teachers feel professionally inadequate, and that this situation damages professional solidarity. The opinion that practice is against the principle of equality in education,

creates a hierarchical structure among teachers, and causes a perception of professional inadequacy in teachers who do not have the title of expert or head teacher is prominent [30%].

UÖ1: *"I think it is a practice that is contrary to the principle of equality in education. The right to take the exam was given to teachers with a certain number of years of seniority, and this led to discrimination among teachers. My colleagues, especially those who could not take the expert teacher exam, considered this an injustice."*

UÖ12: *"There were unpleasant situations among teachers. We heard statements like, 'You are the expert, I am the head teacher.' This discrimination created a negative atmosphere among teachers."*

Stress and loss of self-esteem

Some teachers stated that the career ladder process was a stressful experience and that the exam preparation process negatively affected their self-esteem. Teachers stated that they believed that this process was not a fair method of measuring professional competence and that therefore the exam became a source of stress. It is the opinion that the career ladder application causes anxiety and loss of self-esteem among teachers, that it can create a perception of professional inadequacy in teachers who do not have the title of expert or head teacher, and that the exam process creates great pressure on teachers [6.67%]. It was stated that this stress is reflected not only in professional life but also in the daily lives of teachers.

UÖ21: *"We experienced great stress throughout the exam process and we carried this stress into our homes. We felt that we were not given the value we deserved as teachers. Seeing discrimination in the eyes of a parent such as 'He is a head teacher, she is not an expert' made us feel worthless."*

Concerns about material inequality and loss of rights

Some teachers stated that the career ladder application led to financial inequalities and that certain groups, especially retired teachers, suffered a loss of rights. The view that providing teachers with financial advantages based on exams is unethical and that this situation disrupts the economic balance among colleagues is prominent. When the evaluations regarding the sub-theme of financial inequality and loss of rights are examined, it is seen that the career ladder application increases income differences among teachers and that providing additional payments through exams is not considered fair. In particular, the fact that retired teachers cannot benefit from these rights is not considered an objective practice, and it is argued that making title and salary increases conditional on exams is unethical and that these rights should be granted based on seniority [23.3%].

UÖ17: *"There was no need for the state to organize an exam to provide a salary increase to teachers. It would have been more appropriate to give it without an exam. Creating discrimination between teachers in terms of both economic and status by subjecting them to an exam does not seem ethical to me."*

UÖ27: *"This process has been a great disadvantage for our retired colleagues. They cannot even benefit from their right to compensation. I think this situation should be corrected retroactively."*

Waste of time

Some teachers stated that the preparation process for the career ladder exam creates a significant burden in terms of time. It was stated that while teachers should focus on their primary responsibilities, which are education and teaching processes, the time spent on exam preparation increases their workload. In this context, there are prominent views that the exam preparation process causes restrictions in teachers' professional and personal lives and causes them to move away from the basic duties of the teaching profession due to their intense work tempo [6.67%].

UÖ23: *"We spent a lot of time preparing for the exam. During this period, we experienced restrictions both in school and in our private lives. We had to take time to study, and this made it difficult for us to focus on our main responsibilities as teachers."*

Status differences

It has been stated that status differences based on career steps damage professional unity among teachers. It has been stated that separating teachers who apply the same curriculum and teach similar courses with different titles has negative effects on cooperation and integrity in the educational environment. In addition, opinions have been expressed that status differences based on title do not comply with the principles of professional solidarity and equality. Teachers think that the economic and status-based differences caused by the titles of expert teacher and head teacher weaken professional unity and negatively affect unity and solidarity in education [23.3%].

UÖ7: *"Expressions such as 'head teacher', 'expert teacher' etc. can cause divisions in the staffroom. We didn't have such a situation in our school, but we have heard of such divisions in other schools."*

UÖ8: *"I am against discrimination or othering people in any field or subject. In the end, everyone, be it a expert teacher, a head teacher or a regular teacher, takes the same class, teaches the same subject and applies to the same curriculum. In other words, I do not think it is right to single out teachers with an exam. A different solution can be developed."*

Discussion, Conclusion and Recommendations

According to the findings obtained from the first sub-problem of the quantitative part of the study, the trainings and video lessons received before the exams made very little contribution to professional competencies. In the evaluation made according to gender difference, it was found that the perception level of male participants was higher than that of female participants; this finding is consistent with the research of Tosun and Yengin Sarpkaya (2014). When the branch differences were examined, it was determined that expert teachers had similar views on the contribution of pre-exam trainings to professional competencies and

that branch was not an important factor. In addition, no significant relationship was found between exam scores and trainings, and it was concluded that seniority did not have an effect on participation or benefiting from the training process.

The findings obtained from the second sub-problem of the quantitative part of the study revealed that expert teachers thought that the career ladder exam was inadequate in measuring professional competencies. Male teachers evaluated the exam more positively than female teachers, but no significant difference was found between branches. It was determined that the exam did not sufficiently take into account the professional competencies and branch differences of teachers and therefore was inadequate in correctly evaluating teacher performance. It was concluded that the relationship between the scale and exam scores was weak, and therefore its effect on professional development and competencies was limited. In addition, it was determined that seniority did not play a decisive role in professional experience and educational benefits.

The findings obtained from the third sub-problem of the quantitative part of the study reveal that the title of expert teacher does not provide the expected contribution to the professional competencies of teachers. Participants generally evaluated the career stage exam as "It did not contribute at all", and it was revealed that the title was ineffective on critical competencies such as field knowledge, educational planning, creating a learning environment and measurement-evaluation. In terms of gender, male teachers thought that the title contributed more to their professional competencies, while female teachers were found to have more negative perceptions. No significant difference was found between branches, and it was observed that teachers in different branches approached the title of expert teacher in a similar way.

There is no significant relationship between the expert teacher title evaluation scale scores and exam scores, which reveals that exam scores do not have a significant effect on teachers' title evaluations. No significant relationship was found between teachers' seniority and scale scores, and it was concluded that as teachers gain professional experience, this experience does not make a significant contribution to the competencies measured by the scale, and the effect of seniority on the development of teachers' professional competencies is limited.

In general, it was concluded that the training, exam and title received within the scope of the expert teacher career ladder application did not provide the expected contribution to the professional competence of teachers, that gender was effective in the perceptions regarding the effects of the career ladder, but the branch was not decisive. These findings reveal that the expert teacher career ladder application should be re-evaluated and that alternative methods that will contribute more to the professional development of teachers should be developed. These results are consistent with the findings of Bakioğlu and Banoğlu (2013) and Çın (2014); they support the suggestions that the title is ineffective, functionality should be increased and in-service training should be emphasized.

As a result of the analyses made in the fourth sub-problem of the qualitative part of the research, the main themes were shaped around the titles of "positive thoughts", "negative thoughts", "advantages", "disadvantages" in order to examine the experiences and opinions of expert teachers regarding the teaching career stage application regarding the first sub-problem of the study.

When the positive opinions of teachers are examined, it is seen that the career ladder application contributes to personal development and professional skill increase. In the research conducted by Pınar and Dönel Akgül (2023), teachers generally have a positive opinion that career ladders will contribute to professional development and increase motivation. The result obtained and this result in the literature are consistent with each other. Teachers stated that this application offers innovative learning opportunities and has a positive effect on career development and planning. In particular, it was emphasized that exams create a source of motivation for teachers and keep the continuous learning process alive. However, teachers stated that the application should be supported with more effective, practical and daily life usable training. A similar result was seen in the study titled "Teacher Competencies in Türkiye" by Tuğluk and Kürtmen (2018) in the literature. In the study, it was concluded that considering the competencies with a broader perspective and supporting them with examples could contribute to the professional development of teachers.

When negative opinions are examined, it is seen that teachers criticize the exam as unnecessary and not compatible with the nature of the teaching profession. It has been stated that the exam creates unnecessary stress on teachers and causes injustice. In the research conducted by Çobanoğlu and İlkin (2023), it was concluded that teachers generally think that career steps and the exam related to these steps are unnecessary. This finding is quite consistent with the result. In addition, the content and structure of the exam are inadequate, it should be measured with the experiences of teachers and the opinion that professional knowledge and skills cannot be evaluated with the exam is dominant. When we look at the literature, there are many studies that are consistent with this finding. In the studies of Avcı and Kayıran (2023) and Kurt (2007), it was stated that teachers participate in the exam due to financial concerns and this situation negatively affects professional motivation. Finally, it was shared that the teaching profession requires expertise in itself, class distinctions are not correct and the exam is against the nature of the teaching profession.

According to the research results, the majority of teachers stated that the career ladder application provided financial gain and brought economic relief. In the research of Bakırcı, Özkan and Özdemir (2023), the participants also expressed positive views that the law provided economic benefits. In addition, there are also views that this application increased professional knowledge and skills, refreshed existing information and contributed to obtaining up-to-date information. Positive feedback was obtained that the career ladder application increased teachers' motivation and interest in the profession, and also gained prestige in the society,

parents and students. These results are consistent with the finding in Webb's (2016) research that increasing teacher quality contributed to teachers being self-confident and talented, increasing their reputation and increasing the quality of education. From the findings regarding the teaching career ladder application, it was seen that financial advantages and professional development elements were prominent, while factors such as motivation, increased interest and prestige had a more limited effect. In particular, material advantages that provide economic relief constitute the center of teachers' positive views on career ladder implementation.

When the results regarding disadvantages are examined, some teachers stated that the career ladder application contradicts the principle of equality in education and creates discrimination and professional division among teachers. In addition, there are also views stating that this application causes stress and loss of self-esteem and negatively affects the perception of professional competence among teachers who do not have the title of expert or head teacher. Some teachers argued that the application creates financial inequalities and that it is unethical to make extra payments through exams. The fact that retired teachers cannot benefit from these rights and that titles are given through exams is found to be contrary to the principle of equality among teachers. This situation has led to criticism that teachers' expectations of equality in economic and professional terms are not met. It has also been noted that the time teachers spend during the exam preparation process causes them to drift away from their professional duties and that this process causes restrictions both in school and in their private lives. In addition, it has been stated that status differences based on titles contradict the principles of professional solidarity and equality. Teachers think that despite attending the same curriculum and courses, title differences negatively affect unity and solidarity in education. Quantitative data revealed that the training received by teachers prior to the expert teacher exam had a low impact on their professional competencies. Similarly, qualitative data supported this finding, showing that teachers considered the training to be theory-heavy, disconnected from practice, and unrelated to professional application. These findings align with the convergent parallel design of the mixed-method approach, indicating that both types of data point to similar results.

While the quantitative data indicated that the exam was insufficient in measuring competencies, the qualitative data showed that teachers perceived the exam as "excluding experience and pedagogical skills" and "focused on rote memorization." At this point, the data converged.

Quantitative findings suggested that the title had a limited impact on professional competencies, whereas qualitative data revealed that the title provided prestige among teachers but was inadequate in terms of pedagogical development. This presents a structure in which both types of data not only support but also complement each other.

These results show that the teaching career ladder application provides positive contributions to the professional development and motivation of teachers, but the exam-centered structure of the application should be reconsidered. Opinions frequently emphasize that title assignments should be supported by longer-term and practical training, taking into account the experience and professional seniority of teachers. Although the career ladder application provides various advantages to teachers both financially and professionally, it also leads to significant disadvantages such as inequality, discrimination and stress. These findings show that the application should be re-evaluated by developing a system that observes the principles of justice and equality. Such an approach can increase teachers' professional satisfaction and make the application more acceptable. Berry and Eckert (2012) stated in their research that current policies do not sufficiently reflect the real conditions in retaining effective teachers and that teacher incentive policies should be developed in a holistic way to support a more effective teaching model. This finding is parallel to the results obtained regarding the career ladder application. A holistic approach is needed that supports the professional development of teachers while also taking into account their psychosocial and economic needs.

Based on the findings of this study, which utilized convergent parallel mixed methods design combining both qualitative and quantitative approaches, participants' critical perspectives on the objective assessment of the teaching career steps and their impact on competency and professional growth have led to recommendations for improving and expanding the current evaluation and development framework.

The teaching career ladder system should be restructured in a way that is objective and motivating and contributes to the professional development of teachers. First of all, the removal of the teaching career ladder exam and the adoption of an appointment system based on seniority and experience will allow teachers' long-term experiences to be better evaluated. In addition, teachers' participation in field-specific training and seminars can be encouraged, and their professional development can be supported more effectively.

Reducing the salary differences between teachers and adopting a more equitable compensation policy will increase teachers' professional satisfaction. In this context, objective and reliable evaluation criteria should be established in the process of advancing to career levels, and a system based not only on teachers' exam success but also on their professional knowledge and performance should be developed. Seminars and training programs should be made more field-oriented and content appropriate to teachers' real needs should be presented.

Creating psychological and professional support mechanisms that will help teachers cope with exam stress and professional pressures will both increase the quality of education and increase teachers' motivation. In addition, reflecting the privileges gained during the teaching career, such as titles, salary increases and status, to teachers' post-retirement processes will increase the attractiveness of the profession in the long term.

It is also of great importance to develop a reliable evaluation system for expert teaching. The fact that this system is based on objective criteria will strengthen the perception of justice among teachers and will make career advancement more meaningful. In addition, the effects of education policies on teacher motivation and quality of education should be examined comprehensively, and the consequences of career advancement on teachers' stress levels and job satisfaction should be investigated.

Finally, studies should be conducted on how educational technologies contribute to the career development of teachers and the effects of continuous professional development programs on career stages should be evaluated. Such studies will help to create more conscious and effective policies for the professional development of teachers.

Declaration of Conflicting Interests

The researchers have no personal or financial conflicts of interest with other individuals or institutions related to the research.

Statements of publication ethics

In this study, all ethical rules specified in the Directive on Scientific Research and Publication Ethics of Higher Education Institutions were followed. No unethical behavior under the heading 'Actions Contrary to Scientific Research and Publication Ethics' in the Directive has been applied.

Examples of author contribution statements

The research design and theoretical foundation were primarily developed by Author 1, working closely under the supervision of Author 2. Throughout the entire research process, Author 2 offered ongoing guidance, constructive critiques, and substantive revisions. Both authors jointly analyzed the results and collaborated in drafting the discussion and conclusion sections of the manuscript.

Researchers' contribution rate

The researchers equally shared responsibilities in the design, conduct, and writing of this study.

Ethics Committee Approval Information

Ethical permission was obtained for the research from the Inonu University Ethics Commission with the letter dated 14.12.2023 and numbered E.383618.

REFERENCES

- Avcı, A. and Kayıran, D. (2023). Evaluation of Specialist Teacher Training Program within the Framework of General Competencies of Teaching Profession. *Journal of Interdisciplinary Educational Research*, 7 (14) , 1-11. <https://doi.org/10.57135/jier.1229524>
- Bakırcı, H., Özkan, Y. and Özdemir, C. (2023). Opinions of Classroom Teachers about the Teaching Profession Law. *e-Journal of International Educational Research*, 14(4), 325-343. <https://doi.org/10.19160/e-ijer.1302529>
- Bakıoğlu, A. and Banoğlu, K. (2013). Investigation of Teachers' Views on the Application of Career Steps in Teaching Using Metaphors and Social Network Analysis. *Marmara University Atatürk Faculty of Education Journal of Educational Sciences*, 37(37), 28-55.
- Berry, B. and Eckert, J. (2012). Creating teacher incentives for school excellence and equity. National Education Policy Center, 1-21. <https://nepc.colorado.edu/publication/creating-teacher-incentives>
- Büyükoztürk, Ş., Çokluk, Ö., & Köklü, N. (2020). Chapter 4: probability, standard normal distribution, and standard scores. In Ş. Türlüdür & M. Çetin (Eds.), *Statistics for social sciences* (pp. 53-71). Pegem Academy.
- Büyükoztürk, Ş. (2020). Some statistics used in validity and reliability analyses. In Ş. Türlüdür & M. Çetin (Eds.), *Data analysis handbook for social sciences; statistics, research design, spss applications and interpretation*. (pp. 179-194). Pegem Academy
- Büyükoztürk, Ş. (2020). Multivariate statistics. In Ş. Türlüdür & M. Çetin (Eds.), *Data analysis handbook for social sciences; statistics, research design, spss applications and interpretation*. (pp. 133-154). Pegem Academy.
- Can, A. (2020). Determining the structural properties of a measurement tool. In Ö. Yüksek (Ed.), *Quantitative data analysis in the scientific research process* (pp. 319-337). Pegem Academy.
- Can, A. (2022). *Quantitative data analysis in scientific research process with SPSS*. Ankara: Pegem.
- Cassell, C. ,Buehring, A., Symon, G., Johnson, P. & Bishop, V. (2005).Qualitative management research: a thematic analysis of interviews with stakeholders in the field. Report to ESRC, 1-82.
- Cındık, H., & Akyüz, C. (1998). The structure, problems, and solution proposals of small and medium-sized forest products industry enterprises in the province of Trabzon. *Turkish Journal of Agriculture and Forestry*, 22(1), 7–11.
- Constitutional Court, 27173 (Constitutional Court March 18, 2009). <https://normkararlarbilgibankasi.anayasa.gov.tr/Dosyalar/Kararlar/KararPDF/2008-107-nrm.pdf>
- Crehan, L. (2016). Chapter 2. Approaches Using Incentives. In L. Production (Ed.), *Exploring the impact of career models on teacher motivation* (p. 40-41). International Institute For Educational Planning.
- Creswell, J. W. and Clark, V. L. P. (2018). Chapter 3, core mixed methods designs. In H. Salmon (Ed.), *Designing and conducting mixed methods research* (pp. 120-128). Sage Publications.

- Çın, H. (2014). Evaluation of Science Teachers' Level of Having General Competencies for Teaching Profession (Personal and Professional Values- Professional Development Competencies) - Diyarbakır Province Example [Master's Thesis]. Kahramanmaraş Sütçü İmam University.
- Çobanoğlu, F. and İlkin, A. (2023). Teachers' Views on the Renewed Career Steps Application. *IBAD Journal of Social Sciences*, (14), 155-173. <https://doi.org/10.21733/ibad.1213828>
- Doğan, İ. (2016). Education as a social institution. In D. Kestek (Ed.), *Sociology concepts and problems* (pp. 228-230). Pegem.
- Field, A. (2009). *Discovering Statistics Using SPSS: Third Edition*. London: SAGE Publications, <https://www.researchgate.net>, (08.08.2024)
- Kaplan, I. and Gulcan, M. G. (2020). Examining the views on the formation of teacher career steps: mixed method. *Gazi Educational Sciences Journal (GEBD)*, 6(3), 380-406. <https://doi.org/10.30855/gjes.2020.06.03.007>
- Karagözoğlu, G. (2003). A general view of our teacher training policy in our education system, reflections in education. VII. Symposium on Teacher Training in Contemporary Education Systems, Cumhuriyet University Culture Center, (pp. 10-12). Sivas.
- Kaya, A. (2015). Estimation models in adaptive cluster sampling. *Adıyaman University Journal of Science*, 2(2), 105-119.
- Kurt, M. (2007). The Meanings Attributed to the Teaching Career Steps Application by Teachers and Administrators [Unpublished Master's Thesis]. Nigde University.
- Marshall, C. and Rossman, G. B. (2014). *Designing Qualitative Research*. New York: Sage.
- Ministry of Education (2005). *ORegulation on promotion in teaching career ladder*. <https://www.resmigazete.gov.tr/eskiler/2005/08/20050813-2.htm>
- Ministry of Education (2006). *General Competencies for the Teaching Profession*. https://oygm.meb.gov.tr/meb_iys_dosyalar/2017_12/13161921_YYretmenlik_MesleYi_Genel_YETERLYKLERi_onaylanan.pdf
- Ministry of Education (2017). *General Competencies for the Teaching Profession*. https://oygm.meb.gov.tr/dosyalar/StPrg/Ogretmenlik_Meslegi_Genel_Yeterlikleri.pdf
- Ministry of National Education (2022). Teaching Profession Law. <https://www.resmigazete.gov.tr/eskiler/2022/02/20220214-1.htm>
- Ministry of National Education, (2024). Schools and other institutions. Ministry of National Education. <https://www.meb.gov.tr/baglantilar/okullar/index.php?ILKODU=42&ILCEKODU=1>
- Pinar, MA and Dönel Akgül, G. (2023). Views of science teachers on the career promotion system in teaching. *International Journal of Social Sciences and Education Research*, 9(1), 36-47. <https://doi.org/10.24289/ijsser.1181614>
- Rasberry, M. R. (2022). The dynamics that exist among teacher preparation programs, student achievement, and teacher performance [Unpublished doctoral thesis]. University Of Houston.
- Seidman, I. (2006). *Interviewing as qualitative research: A guide for researchers in education and the social sciences*. Teachers College Press.
- Tosun, İ. and Yengin Sarpkaya, P. (2014). Teachers' views on career stages. *Turkish Studies (Electronic)*, 9(5), 1971-1985. <http://dx.doi.org/10.7827/TurkishStudies.6713>
- Tuğluk, MN and Kürtmen, S. (2018). Teacher competencies in Türkiye. *Adıyaman University Journal of Social Sciences Institute*, 10(30), 810-841. <https://doi.org/10.14520/adyusbd.480236>
- Tutar, H. (2022). Validity and reliability in qualitative research: a model proposal. *Anadolu University Journal of Social Sciences*, 22(2), 117-140.
- Tutar, H. and Erdem, AT (2022). *Sampling techniques*. In *Scientific research methods and spss applications with examples* (pp. 241-277). Seçkin Publishing.
- Urfalı, P. (2008). Views of Teachers Working in Primary and Secondary Education Institutions on the Promotion System in Teaching Career Steps- Eskişehir Province Example [Master's Thesis]. Anadolu University.
- Üstüner, M. (2002). *Scientific foundations of education*. On education. E. Toprakçı (Ed.). Ankara: Ütopya Publishing House.
- Webb, MA (2016). The investigation of interns, teachers, and principals perceptions of preparedness in elementary education [Unpublished doctoral dissertation]. Auburn University.



| Research Article / Araştırma Makalesi |

When Metacognition Works—And When It Doesn't: Investigating the CRIME Technique in Middle School Mathematics

Üstbilişsel Yöntem Herkese Uygun mu? CRIME Tekniğinin Matematik Dersinde 7. Sınıf Öğrencilerine Etkisi¹

Görkem KILIÇ², Mesut ÖZTÜRK³

Keywords

Metacognition
Problem-solving
Ratio
Proportion
Percentages

Anahtar Kelimeler

Üstbiliş
Problem çözme
Oran
Orantı
Yüzde

Received/Başvuru Tarihi
03.04.2025

Accepted / Kabul Tarihi
10.10.2025

Abstract

Purpose: This research aimed to examine the effect of teaching ratio-proportion and percentages on subjects of 7th-grade students using the CRIME technique, one of the metacognitive teaching methods.

Methodology: The case study model, one of the qualitative research methods, was used in the study. Qualitative studies examining the learning process was determined, and the researcher prepared five semi-structured interview items for the interview form. The study group consisted of six students studying in the 7th grade at a public school in a province in the northeast of Turkey. Purposeful sampling method was used to determine the participants of the study. For this purpose, participants were determined from the 24-person student group to which the application was made using the maximum diversity sampling method. The selected students were identified by asking the questions in the interview form.

Findings: The result of the study show that the metacognitive CRIME technique increased motivation, problem-solving, course success, and metacognitive skills. The study determined that the metacognitive CRIME technique improved metacognitive monitoring and evaluation skills in high-achieving students.

Highlights: At the same time, it did not show any change in the metacognitive skills of students with medium and low levels of success, which is an original result. Unique result of the study is that the metacognitive CRIME technique was effective in the problem-solving skills of medium-achieving students. At the same time, there was no change in the problem-solving skills of low-achieving students.

Öz

Çalışmanın amacı: Bu araştırmanın amacı, 7. sınıf öğrencilerine meta-bilişsel öğretim yöntemlerinden biri olan CRIME tekniği kullanarak oran-orantı ve yüzdelerin öğretiminin dersler üzerindeki etkisini incelemektir.

Yöntem: Çalışmada nitel araştırma yöntemlerinden durum çalışması modeli kullanılmıştır. Öğrenme sürecini inceleyen nitel çalışmalar belirlenmiş ve araştırmacı görüşme formu için beş adet yarı yapılandırılmış görüşme sorusu hazırlamıştır. Çalışmanın çalışma grubunu Türkiye'nin kuzeydoğusundaki bir ilde bulunan bir devlet okulunda 7. sınıfta öğrenim gören altı öğrenci oluşturmuştur. Çalışmanın katılımcılarını belirlemek için amaçlı örnekleme yöntemi kullanılmıştır. Bu amaçla uygulama yapılan 24 kişilik öğrenci grubundan maksimum çeşitlilik örnekleme yöntemi kullanılarak katılımcılar belirlenmiştir. Seçilen öğrenciler görüşme formunda yer alan sorular sorularak belirlenmiştir.

Bulgular: Çalışmanın sonucunda, üstbilişsel CRIME tekniğinin motivasyonu, problem çözme, ders başarısını ve üstbilişsel becerileri artırdığını göstermektedir. Çalışma, üstbilişsel CRIME tekniğinin yüksek başarı gösteren öğrencilerde üstbilişsel izleme ve değerlendirme becerilerini geliştirdiğini belirlemiştir.

Önemli Vurgular: Orta ve düşük başarı seviyelerine sahip öğrencilerin üstbilişsel becerilerinde herhangi bir değişiklik göstermemiştir ki bu orijinal bir sonuçtur. Çalışmanın özgün sonucu, üstbilişsel CRIME tekniğinin orta düzeyde başarılı öğrencilerin problem çözme becerilerinde etkili olduğudur. Aynı zamanda, düşük düzeyde başarılı öğrencilerin problem çözme becerilerinde bir değişiklik olmamıştır.

¹ This study was supported by Bayburt University Scientific Research Project Coordinatorship under Grant Number 2024/69001-01-01. The authors thank Bayburt University Scientific Research Project Coordinatorship for their support.

² Turkish Ministry of National Education, Bayburt, TÜRKİYE; <https://orcid.org/0000-0001-6289-1847>

³ **Corresponded Author**, Bayburt University, Faculty of Education, Department of Mathematics and Science Education, Bayburt, TÜRKİYE; <https://orcid.org/0000-0002-2163-3769>

INTRODUCTION

"A few years ago, I taught my students the subject of ratio-proportion and percentages using eggs scrambled with tomatoes and green pepper recipe. The students calculated the ingredients for scrambling eggs with tomatoes and green pepper. Then, they made different calculations using pilaf and cake recipes. Afterward, the students took a national exam, and after the exam, one of my students came to me and said, "Teacher, you gave eggs scrambled with tomatoes and green pepper recipes, pilaf recipes, and cake recipes in class. However, they did not ask any of them about the apple cookie recipe in the exam; they asked about it. I could not answer the question because you did not teach us the apple cookie recipe." Upon this statement from the student, I realized that the daily life example I used to teach ratio-proportion did not contribute to students' generalization and thinking skills. As a result of this experience and my research that formed the basis for this research, I realized that metacognitive training could be effective in students' establishing the connection between mathematics and daily life."

Metacognitive teaching can provide meaningful learning by making students aware of the learning process (Erdem & Öztürk, 2023; Okumuş & Öztürk, 2024). Erdem and Öztürk (2023) determined that students are more willing to learn when metacognitive teaching is supported by technology. Due to the above reasons, the metacognitive CRIME technique was used to teach ratio-proportion and percentages in this study. The metacognitive CRIME technique is a teaching approach to ensure students understand mathematical concepts and increase their metacognitive awareness. The metacognitive CRIME technique, developed in a technology-supported environment, helps students structure their mathematical thinking processes and develop metacognitive strategies in the problem-solving process (Teong, 2003).

Theoretical Framework

Metacognition

Metacognition is defined as thinking about thinking, and the first important studies on this subject were shaped by the studies of Flavell (1979). Flavell emphasized having knowledge about cognitive processes and being able to manage these processes in his studies. In other words, metacognition is the individual's awareness of his own thinking process and being able to control it. In this context, the individual can determine learning strategies, evaluate which methods to use in which situations, and make the necessary arrangements to make the learning process more effective (Prins, Veenman, & Elshout, 2006). Based on this understanding, two aspects of metacognition are distinguished: (1) knowledge and (2) regulation of cognition. The components of metacognition are shown in Figure 1.

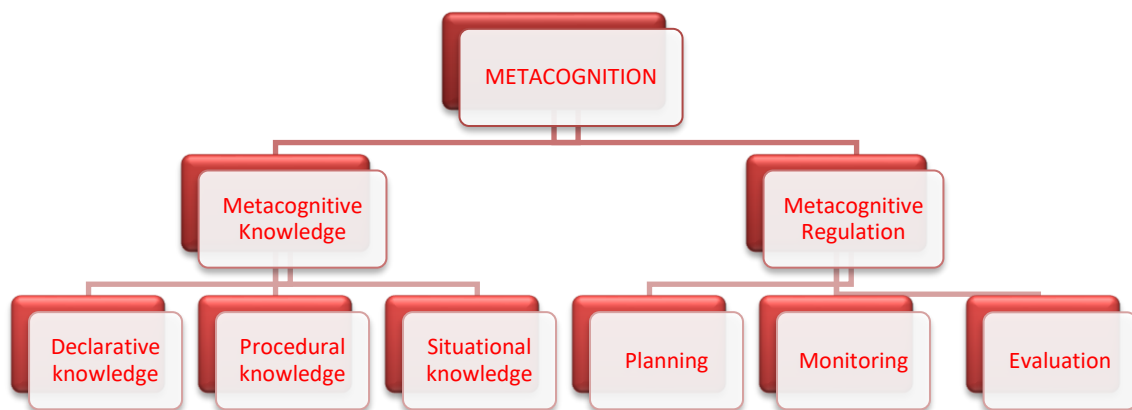


Figure 1. Components of Metacognition

While metacognitive regulation is examined under three headings: planning, monitoring, and evaluation, metacognitive knowledge is examined under three headings: declarative knowledge, procedural knowledge, and situational knowledge.

Knowledge of cognition includes declarative knowledge about strategy, task, and person (Baumanns & Rott, 2022). Strategic knowledge refers to strategies (e.g., in problem-solving) and when to apply them. Task knowledge refers to knowledge about the different levels of difficulty of tasks (e.g., in problem-solving) and the different strategies required to solve them. Person knowledge includes knowledge about one's own strengths and weaknesses (Baumanns & Rott, 2022).

Refers to the procedural knowledge related to the processes that coordinate cognition. It includes planning, monitoring, and evaluation activities (Baumanns & Rott, 2022). Planning refers to determining a goal related to the current effort, activating previous content knowledge, and activating metacognitive knowledge, such as knowledge about specific tasks and how to solve them. For example, when a student prepares for a math exam, the student determines which topics the students will study and decides how much time he will allocate to each topic. In this process, he plans which topics require more repetition by recalling his previous knowledge. Monitoring refers to metacognitive awareness and monitoring of cognition; for example, when an individual realizes that he understands a problem-solving task, it is monitoring.

Another example is a student who thinks he/she understands a text he reads, but when he/she asks himself a few questions, he/she realizes that he/she does not fully understand some concepts. This situation shows that the student monitors the learning

process and identifies his deficiencies. Evaluation refers to the activities of reflecting and judging one's own performance and results in the form of a result. For example, A student who looks at the questions he got wrong after an exam, analyzes which topics he made mistakes on, and decides that he needs to focus more on these topics in the following study period can be considered an expression of evaluation. These activities are not entirely different, and perspectives on their overlap may exist.

Metacognitive training

Many teaching techniques based on metacognition have been developed. Some of them are SQ3R, MURDER, IMPROVE and CRIME.

- SQ3R (Survey, Question, Read, Recite, Review): It consists of the stages of browsing, asking questions, reading, repeating, and reviewing (McCormick & Cooper, 1991).
- MURDER: Metacognitive technique that allows students to access information more consciously and effectively during the reading process (Hythecker, Dansereau, & Rocklin, 1988).
- IMPROVE: It consists of introduction, metacognitive questioning, application, review, gaining expertise, verification, and enrichment. It is a technique based on metacognitive questioning and aims to develop students' thinking processes (Mevarech & Kramarski, 1997).
- CRIME: It includes the steps of careful reading, choosing the appropriate strategy, implementing the strategy, monitoring, and evaluating. CRIME is a problem-solving-focused technique and enables students to use solution strategies consciously (Teong, 2003).

In this study, an application based on the metacognitive CRIME technique was made, and only this technique was explained in detail.

Teaching based on the metacognitive CRIME technique

In this study, the CRIME technique, one of the metacognitive training techniques, was used. The stages of the CRIME technique are shown in Table 1.

Table 1. CRIME strategies

Strategies	Explanations
Careful reading	At this stage, the student reads the problem, determines the connection of the problem to previous problems, determines whether he/she fully understands the problem, and writes down what is given and asked in the problem.
Recall possible strategies	The student chooses or remembers the strategy that he/she thinks will be appropriate for solving the problem based on his/her previous solutions. The student may be drawing a figure, breaking the problem into pieces and solving it part by part, or simplifying it.
Implement strategy	The student must implement the strategy he/she has determined as planned.
Monitor	The individual asking himself questions should monitor whether the strategy is implemented correctly.
Evaluation	The result is checked operationally, and its logicity is discussed.

Table 1 shows that the metacognitive CRIME technique consists of acrostic letters. The CRIME technique strengthens students' cognitive processes and contributes to developing their thinking, analysis, and reasoning skills. It helps students to address complex problems more consciously by providing a systematic and structured approach to the problem-solving process (Teong, 2003). This technique encourages students to analyze data by supporting their evidence-based thinking, developing their interpretation skills, and building their solution processes on solid foundations. Students can think and question more deeply by avoiding quick and superficial guesses (Kwang, 2000).

Adopting a student-centered approach, the CRIME technique moves away from traditional rote learning methods and allows students to participate in the learning process actively. It contributes to developing individuals who can think independently and manage their learning processes, enabling students to develop analytical and critical thinking skills (Kwang, 2000). At the same time, the CRIME technique increases students' self-confidence by helping them trust their own thoughts and express themselves while producing solutions. As they become successful in problem-solving, it helps students develop a positive attitude toward mathematics. Then, the CRIME technique contributes to strengthening cooperation and discussion culture. It encourages students to share information and evaluate different perspectives in group work. Creating an environment of interaction and discussion in the classroom allows students to learn from each other and defend their own ideas (Kwang, 2000). The CRIME technique supports academic success and helps students apply the information they learn to real-life problems. It enables them to use their problem-solving skills in different areas of life by making them realize that mathematics is an important tool in daily life (Teong, 2003).

Literature Review

The literature has examined and determined that studies on metacognition mainly focus on problem-solving. However, although there are applications aimed at teaching mathematics to students through metacognitive strategies, they are few (Aydın & Öztürk, 2024). It is emphasized that research on metacognitive training is important in developing metacognitive skills in students (Ares Ferreirós et al., 2024; Gandolfi et al., 2024; Okumuş & Öztürk, 2024). In this direction, many studies have been conducted to develop metacognitive training techniques at international and national levels (Hythecker et al., 1988; Erdem & Öztürk, 2023; Kwang, 2000; McCormick & Cooper, 1991; Mevarech & Kramarski, 1997; Okumuş & Öztürk, 2024). However, studies examining the effectiveness of these techniques are minimal, especially at the national level (Öztürk, 2021; Şahinkaya et al., 2022). In this context, it can be said that this study, which will evaluate the effect of metacognitive training in mathematics teaching, is important. Mevarech and Kramarski (1997) presented IMPROVE techniques based on metacognitive questioning, and Kwang (2000) presented CRIME techniques for metacognitive problem-solving. In studies conducted in Turkey, metacognitive planning and monitoring-based teaching techniques were emphasized (Okumuş & Öztürk, 2024). Since today's education systems are based on problem-solving and questioning, teaching was done with the metacognitive CRIME technique in this study.

Metacognition intervention research and compilation studies have also emerged in recent years. For example, Aydın and Öztürk (2024) conducted a meta-analysis study to express the contributions of studies examining the effects of metacognitive training on students' mathematical achievement and problem-solving skills. The study included 12 experimental studies between the years 2000-2023. They stated that students who received metacognitive training teaching were more successful in mathematical achievement and problem-solving skills. Eberhart et al. (2025) conducted a meta-analysis study to examine whether metacognition interventions were effective in young children. They conducted the study with preschool and primary school students. As a result of the study, they determined that metacognitive interventions increased children's self-efficacy and that the positive effect could last for a long time. The study shows how teachers and researchers collaborate to develop intervention programs together. In addition, it was found that young students benefit from metacognitive interventions in various ways, supporting the evidence and expanding their knowledge in this field.

The Present Study

This study used the metacognitive CRIME technique to examine the teaching process of ratio-proportion and percentages to 7th-grade students. To achieve this aim, the following research questions were sought:

1. Metacognitive CRIME technique in teaching ratio-proportion and percentages in 7th grade?
2. The metacognitive CRIME technique in problem-solving in teaching ratio-proportion and percentages in 7th grade?
3. Metacognitive CRIME technique in teaching ratio-proportion and percentages in 7th grades on course success?
4. Metacognitive monitoring in teaching ratio-proportion and percentages in 7th grades using the metacognitive CRIME technique?
5. Metacognitive CRIME technique in metacognitive assessment in teaching ratio-proportion and percentages in 7th grade?

METHOD

Research Model

A case study model, one of the qualitative research methods, was used in the study. A case study is a detailed examination of a specific section of one or more events with clearly defined boundaries (Merriam, 2013). This study conducted a detailed examination of a limited situation since the aim was to teach the subject of ratio-proportion and percentages with the metacognitive CRIME technique. For this reason, the case study model was used.

Merriam (2013) divides the case study into three types: historical and observational case studies, internal case studies, and multiple case studies. An instrumental case study was used. An instrumental case study is a detailed examination of a single event to find the underlying facts and generalizations (Merriam, 2013). In this study, an instrumental case study was conducted because the teaching process of the subject of ratio-proportion and percentages was examined as a single case with the metacognitive CRIME technique.

Participants

The study group consisted of 7th-grade students in a public school in northeast Turkey. First, 24 students were studied, then six students were selected from this group, and the primary data were collected from these students. The maximum variation sampling method was used to determine the study participants. In order to provide maximum variation, the student's willingness to participate in the course, gender, and the students' most recent mathematics course score averages were used. To do this, the students were first ranked from the highest to the lowest according to their mathematics course success scores. Then, the eight students with the highest scores were determined as high, the eight with the lowest were determined as low, and the others were medium successful. The students in each group were divided into girls and boys. Three girls and three boys from these students who were willing to participate in the course were determined to be the study participants. Direct quotations were made from

the participants' opinions in the study. Code names were used for the participants in this process. Sij (i: success status [H: High successful, M: Medium successful, L: Low successful]; j: gender [G: Girl, B: Boy]) coding was used in the coding of the participants.

Data Collection Tools

The study data were collected using a semi-structured interview form and a semi-structured observation form.

Semi-structured interview form

The researchers developed the semi-structured observation form used in the study. Firstly, the literature reviewed. Qualitative studies examining the learning process were determined, and the researchers prepared five semi-structured interview form items for the interview form. The form was presented to three mathematics education experts (one professor and two associate professors), and they were asked to evaluate the questions' suitability. One expert stated that two questions could be understood similarly (questions aimed at metacognitive monitoring and evaluation) and suggested corrections. After the necessary corrections were made, the form was presented to a Turkish teacher, and he was asked to evaluate the understandability of the questions. The Turkish teacher stated that the questions were understandable. The form was finally applied to two students not included in the study sample as a pilot application and a pilot study was conducted. The pilot study determined that the students could understand the questions and give sufficient answers. Thus, the form was finalized. An example of the questions in the form is as follows: "Did the activities (applications) done in this course affect your problem-solving skills? Why?"

Semi-structured observation form

The researcher developed the semi-structured observation form used in the study. Merriam (2013) stated that the person making the observation should determine specific characteristics and start the observation. In this context, specific themes were created in the study, and observations were made based on these themes. For this purpose, observations were created according to motivation, problem-solving, course success, metacognitive monitoring, and metacognitive evaluation. The researcher noted the observations in the gaps during and at the lesson's end.

Implementation Process

The study lasted five weeks, and thirty-five lesson hours were applied. The study applications were carried out in the school where the researcher worked. A total of 11 activities were carried out during the study process. At the end of the implementations, interviews were conducted with the students. A sample activity of the study process is presented below.

Careful reading

First, the problem in the activity is presented to the students. Students are asked to read the problem individually. Then, one student reads the problem out loud. The problem in the textbook is given below.

The table below shows the milk used in cheese making and the amount of cheese produced from this milk.

Table 2. Amounts of milk used and cheese produced from milk

Amount of Milk Used (liters)	...	10	...	35	...	60
Amount of Cheese Produced (kg)	1	2	5	7	9	12

In the table above, determine the amount of milk used according to the cheese produced.

- How many liters of milk make 1 kg of cheese?
- How many liters of milk are needed to obtain 20 kg of cheese?
- How many kilograms of cheese can be obtained from 55 liters of milk?
- What is the relationship between the amount of milk used and the amount of cheese obtained? Explain.

Before starting the solution, students are asked to write down what they understand from the problem. Then, they are asked to discuss what they understand from the problem in a small group discussion. During the discussions, students are guided to understand the relationship between the data. During this process, students are asked questions such as: "What *information is in this table?*" and "How *does the amount of cheese change as the amount of milk increases?*" These questions help students clarify the parts of the question that they do not understand.

Recall possible strategies

Students are encouraged to explain the relationship between the amount of milk and cheese in their own words. For example, "Is there a ratio between the amount of cheese obtained and the amount of milk used?" can be asked to make students think. In addition, they are associated with their previous knowledge to remember the concept of ratio. For this, the concept of ratio at the 6th-grade level is reminded, and they are encouraged to associate the concept of ratio at the 6th-grade level with the new concepts they will learn. At this stage, students are given small and large group discussions to choose the correct strategy. After the discussions, students who change their solution strategy are asked to explain why they changed their strategy.

Implement strategy

Students try to complete the missing data based on the table. For example, when solving how many kilograms of cheese can be obtained from 10 liters of milk, students make calculations using the rules of ratio and proportion. Then, they share their solutions in class and discuss different ways of thinking.

Metacognitive monitoring

Students are asked to review the steps they followed in the solution process. They are asked to monitor their own thinking processes with questions such as *“How did you solve the problem?”* and *“What strategy did you use to reach the result?”* If an incorrect conclusion is reached, they are guided to find the source of the error.

Metacognitive evaluation

In the last stage, students interpret the table and answer, *“What is the relationship between the amount of milk used and the amount of cheese produced?”* The problem is verified with the students. At this stage, students are asked to evaluate their own solutions and different solutions, if any, and to think about which solution is more effective. Then, it is discussed whether the problem can be solved with different solutions, and the solutions are tried. In addition, students are asked to evaluate their own performance in the learning process.

Data Analysis

In the analysis of the study's data, previous qualitative studies on teaching effectiveness with the CRIME technique were examined, and a data analysis framework was created using the findings of these studies. Okumuş and Öztürk (2024) categorized students' statements into metacognitive planning, metacognitive monitoring, metacognitive prediction, metacognitive awareness, and motivation in their action research on metacognitive planning and monitoring-based teaching. Teong (2003) revealed that metacognitive training improved students' problem-solving skills. Öztürk (2021) determined that metacognition positively affected students' algebra success and problem-solving as a qualitative finding in his mixed-method research on metacognitive training. The findings of these studies addressed the categories of motivation, problem-solving, course success, metacognitive monitoring, and metacognitive evaluation. The categories and their explanations are given in Table 3.

Table 3. Categories used in the study and their descriptions

Categories	Verbal expressions
Motivation	✓ The connection with daily life is interesting. ✓ Active participation is interesting. ✓ The reward is interesting.
Problem-solving	✓ Supports different solutions. ✓ It enables creative thinking.
Academic success	✓ Can identify deficiencies. ✓ Controls the stages.
Metacognitive monitoring	✓ Identifies learning deficiencies in the learning process. ✓ Identifies learning deficiencies by asking questions to oneself while learning.
Metacognitive evaluation	✓ Identifies learning deficiencies at the end of the learning process. ✓ At the end of the learning process, the student asks himself/herself questions and determines his/her learning deficiencies.

The researcher first coded the data collected with the help of interview forms in the study, as shown in Table 3. Then, the codings were examined by the advisor faculty member, and the codes determined to be inappropriate were removed from the code matrix. The data obtained from the interview were presented according to the remaining codes. At the same time, teacher observations were presented by the categories in Table 3.

Validity and Reliability

In order to ensure the validity of the research, the research process was carried out under the supervision of the advisor faculty member. The research process was planned with the advisor faculty member, and an independent researcher checked the plans for all activities. The advisor faculty member observed at least one class hour each week. At the same time, the researcher's coding was checked by the advisor faculty member.

FINDINGS

The study reported on motivation, problem-solving, course success, metacognitive monitoring, and evaluation.

Findings Regarding Motivation

The opinions obtained from the participants were first collected in the motivation category. Direct quotes from student opinions on motivation are shown in Table 4.

Table 4. Student views on motivation

Motivation	Opinions
S _{HG}	Yes, it caught my attention. After these activities, I can understand the problems better. In the past, when I could not solve the problems, I would immediately give up; now, I try to solve them by thinking more... I also participated more in the lesson... The activities we did helped me focus on the lesson. I also easily associated the problems we solved with daily life.
S _{HB}	Yes, these lessons caught my attention. I liked the rewards my teacher gave after correct answers at each stage. The rewards increased my interest in the lesson; our teacher should give us rewards in every lesson [Student laughed] ... Our teacher encouraged us to come to class prepared. He also rewarded those who came to class prepared; we are delighted when we receive rewards.
S _{MG}	It caught my attention. At first, I had a hard time because it was an activity I was not used to, so I felt like I was behind my friends. However, over time, our activities started catching my attention. Our teacher, especially, asked us to read again and gave us time to think, which helped me understand better.
S _{MB}	The activity caught my attention... While learning the subject, I learned to write understandably and concisely. When solving a problem, my teacher asked me to summarize it and underline the important parts, making it easier for me to learn... Then, reminding me of the previous topics we solved to understand the problem made it easier for me to understand. For this reason, I tried harder to solve it. However, there were some parts that I had difficulty with during the solution stages.
S _{LG}	It did not interest me. It did not interest me because I do not like very active participation. The teacher wanted us to think all the time. I cannot do the math; the teacher wants us to think... I do not want to talk in class; the teacher tries to make us talk. Also, it would be a waste of time if everyone participated in the lesson.
S _{LB}	The activity did not interest me. Not everyone progressed at the same pace; I did not receive awards. The teacher gave pluses to my friends but not to me. She told me to study and think but did not give me pluses. She should not have given pluses to them either. So, when my friends got pluses, and I did not, attending the lesson became boring.

Table 4 shows that the motivation of the participants with high and medium levels of success increased during the learning process with the metacognitive CRIME technique. However, the table pointed out that the students with low success could not be motivated in these activities. The male student with low success indicated that his motivation was low due to not receiving symbolic reinforcement. The female student with low success emphasized that her motivation did not increase because she was not satisfied with active participation.

In the observations made by the researcher, it was determined that the student's interest in the activities carried out with metacognitive CRIME increased and that symbolic reinforcements ensured the active participation of the students. The researcher observed that the students understood the subject better and could focus their attention, especially on the explanations or solved questions associated with daily life. The researcher determined in his observations that the slow learners were reluctant to learn because the individual learning speeds of the students were not the same.

Findings for Problem Solving

The second category, based on the participants' opinions, is problem-solving. Direct quotes from students' opinions on problem-solving are shown in Table 5.

Table 5. Student views on problem solving

Problem-Solving	Opinions
S _{HG}	Yes, it affected my problem-solving. I learned how to understand the subject in summary while solving a problem. I saw what we could do to understand the problem with our activities. I learned what I can do to understand the problem when I do not understand it. It helped me learn different solutions.

S_{HB}	I think this activity has positively affected my problem-solving... I used to focus only on finding the solution to the problem, but now I try to understand the process... I think I am experiencing the problem now [Student laughed]... Because I can solve questions faster, I have learned to think more creatively in the areas where I have difficulty.
S_{MG}	Yes, it did. Even though some questions were difficult, I learned to understand and solve problems easily thanks to the activity... In the past, when there were questions, I had difficulty solving, I would immediately quit. Now, I try to look at it differently. I try to think more thoroughly to understand the problem. I try to make sure I understand it. I even reread the problem most of the time.
S_{MB}	Yes, the activity affected my problem-solving a little. For example, in the past, I did not bother with understanding the problem and would try to solve it immediately. Now, I try to understand the problem. If necessary, I write down what is given and requested and then try to solve it... Thanks to this activity, I learned different solutions for the points I struggled with.
S_{LG}	I have not improved much in problem-solving. I still have difficulty in doing some steps. Some of my other friends in the class can solve problems, but I cannot. That is why I sometimes feel bad... I am not good at problem-solving.
S_{LB}	It does not affect my problem-solving; doing calculations is difficult. I cannot solve problems because I cannot do calculations. The teacher also thinks I cannot solve problems [Student laughed] ... My teacher finds math difficult for me. I cannot do the math. If I do the math, I can solve problems; I do not know.

Table 5 shows that the participants with high and medium levels of success thought that their problem-solving skills improved during the learning process with the metacognitive CRIME technique. However, it was determined that the students with low success thought there was no improvement in their problem-solving skills.

In the observations made by the researcher, it was observed that the students performed systematic operations step by step while learning with the metacognitive CRIME technique and solving problems. The researcher determined in his observations that the students could look at the problem differently when they saw different solution methods. At the same time, it was determined in the researcher's observations that the students who had difficulty determining the strategy could use metacognitive questions.

Findings Regarding Course Success

Another category based on the opinions obtained from the participants is course success. Direct quotes from student opinions on course success are shown in Table 6.

Table 6. Student views on course success

Course Success	Opinions
S_{HG}	Yes, I was able to determine my learning deficiencies. Since I understood the subject well, I immediately noticed my deficiencies and tried to eliminate them. I realized I could determine my learning deficiencies, especially by asking myself questions... I also found my deficiencies by asking myself questions while learning the subject. I thought more about the deficiencies I determined, and when I could not understand, I asked you (the teacher). I think I was able to learn the subject better this way.
S_{HB}	Thanks to this activity, I could identify my learning deficiencies more easily. Because we went step by step in our activity, I could easily control each step. I focused on the step I was missing and corrected it... Sometimes, I found it by thinking on my own, sometimes I consulted my friend, and sometimes I asked the teacher. In this way, I learned the subject better by determining the subjects I was missing and the parts I could not learn.
S_{MG}	My success in the lesson has increased. Especially making connections with daily life helped me understand the subject better... Because the problems we solved in the lesson were problems related to our daily lives... However, there are still some parts that I need to improve. For example, I do not know what I need to do to fill my gaps in knowledge... I understand when you (teacher) teach, but I think I cannot learn when I study alone.
S_{MB}	I was able to identify some of my learning deficiencies. Especially in class, when you (the teacher) asked us to ask ourselves questions, I answered some of the questions I asked but could not answer others. Then, I realized that there were areas where I could not learn. However, I could not do this throughout the class... While I quickly realized my deficiencies in some subjects, I insistently repeated the same mistake in some subjects... When I did not realize that I was making a mistake, sometimes my group mates and you told me I was making a mistake. Therefore, I was only able to identify some of them.

S_{LG}	I had a hard time identifying my learning gaps. I usually have difficulty understanding what I lack because I lack basic information. My lack of knowledge in basic subjects prevents me from learning. Because when you say something, sometimes I understand it, but when I do what you say, I realize I do not know the information I need to use. When this happens, I cannot do the part you said.
S_{LB}	I could not determine my shortcomings in this activity because I had difficulty doing the steps. I tried to understand what you said, but I still could not do it... I could not understand at which stage I made a mistake or did something incomplete.

Table 6 shows that the participants with high levels of success thought that their course success increased during the learning process with the metacognitive CRIME technique. It was determined that the participants with medium and low success levels thought there was no improvement in their course success.

In the observation made by the researcher, it was determined that seeing different solutions during the lesson caused the students to question their own solutions. Therefore, they studied more and did more research. It was observed that the metacognitive CRIME technique increased the students' success by facilitating their learning of the subject of ratio and proportion. It was determined that it encouraged the participation of students with low levels of subject or question solutions associated with daily life.

Findings Regarding Metacognitive Monitoring

Another category based on the views obtained from the participants is metacognitive monitoring. Direct quotes from the student views on metacognitive monitoring are shown in Table 7.

Table 7. Student' opinions on metacognitive monitoring

Metacognitive Monitoring	Opinions
S_{HG}	In this course, I determined my learning deficiencies while learning. When I could not do what you said while learning the course, I checked its accuracy by asking you (the teacher) or my friends. The solutions I found were usually correct... While learning, I sometimes tried to solve the problems I wrote myself. With this application, it was easier for me to find the places where I got stuck in the solution stages... In other words, while learning the topics I could not understand, I noticed my deficiencies and tried to correct them.
S_{HB}	Yes, I could determine the parts I could not learn while learning in class. Our teacher asked us to summarize what we learned. I started to write my learning in summary form, and at this stage, I could not fill in some parts. I realized that I was lacking by asking myself questions in this area. After determining my deficiencies, I sometimes asked my teacher, and sometimes I moved forward by looking at the book. However, seeing my deficiencies allowed me to learn the subject in more detail.
S_{MG}	I partially identified my learning deficiency. I saw my deficiencies or mistakes in some subjects, but I could not always see my deficiencies or mistakes. My friends or my teacher usually told me where I made mistakes... Of course, there were a few times when I found my own mistakes, but they were rare.
S_{MB}	The lesson increased my difficulty. Going step by step and actively participating in the lesson affected my success. However, I had difficulty in some parts. For example, when choosing a strategy, I could not find which one to choose right away. When I talked to my friends, I understood from what they said that I was making a mistake.

Table 7 showed no improvement in the metacognitive monitoring skills of low-achieving students. It pointed out that the metacognitive monitoring skills of the high-achieving participants have improved during the learning process with the metacognitive CRIME technique. It emphasized that the metacognitive monitoring skills of the medium-achieving participants have partially improved.

In the researcher's observations, it was observed that the metacognitive CRIME technique made the students aware of what they knew (metacognitive awareness) and that they could make inferences about what they did not know. Similarly, the researcher determined that students lacking knowledge could not use clues, and their metacognitive awareness was low.

Findings Regarding Metacognitive Evaluation

The last category, based on the opinions obtained from the participants, is metacognitive assessment. Direct quotes from the student opinions on metacognitive assessment are shown in Table 8.

Table 8. Student' opinions on metacognitive evaluation

Metacognitive Assessment	Opinions
S _{HG}	Thanks to these activities, I saw my strengths and weaknesses. I need to see my strengths and weaknesses. This way, I improved my weaknesses. Similarly, since I asked myself questions, I easily covered my shortcomings. Asking myself questions allowed me to see what I knew and did not know at the end of the lesson and what I learned incompletely.
S _{HB}	I realized my own shortcomings in these activities. Thus, I was directed to study more. Thanks to this activity, I learned many strategies. At the end of the lesson, I asked myself questions, saw whether my performance was sufficient, and evaluated whether I reached the correct result. Thus, I became aware of my weaknesses, incomplete learning, and the places I could not learn... My teacher, I knew many things but was unaware of what I knew [Student laughed].
S _{MG}	At the end of this lesson, I partially determined my shortcomings. At the end of the lesson, I wanted to write what I learned and what I did not learn, but I had difficulty writing what I learned in this section and what I did not learn. I am aware of my shortcomings in the activity, but I could not find all of them.
S _{MB}	At the end of the lesson, I determined my deficiencies. Before the activity, I did not realize I had deficiencies, but thanks to this activity, I could go back to the steps and determine them a little. Since I was able to determine my deficiencies, I now know the parts I could not learn. For example, when I start the next lesson or during the break, I know the questions I will ask you and the parts I did not understand. I did not know before... When I looked at my friends while writing what we learned at the end of the lesson, some friends wrote more than me. I realized that I could not learn some of the things they wrote.

Table 8 showed no improvement in the metacognitive evaluation skills of low-achieving students. It pointed out that the metacognitive evaluation skills of the high-achieving participants improved during the learning process with the metacognitive CRIME technique. It emphasized that the metacognitive monitoring skills of the medium-achieving participants partially improved.

In the observations, the students realized their learning by self-evaluating at the end of the learning process. At the same time, it was determined that the students with high success did not notice the errors or deficiencies in the solution stages before the application process. However, after the application, they checked and questioned every process step. The observations determined that the students realized the learning process and understood how they could learn. It was observed that some of the students evaluated the process, while others had difficulties in evaluation.

DISCUSSION

According to the study's findings, the activities implemented with metacognitive CRIME attract the attention of highly successful students, encourage them to think, and enable them to establish a relationship with daily life. Okumuş and Öztürk (2024) stated that metacognitive planning and monitoring-based teaching increase students' motivation, provide motivation, improve attention control, and create positive emotions. Sungur and Senler (2009) also determined that metacognitive training increases students' motivation. Dignath et al. (2008) conducted a meta-analysis showing that metacognitive strategy training positively affects learning outcomes, strategy use, and motivation. Motivation is one of the most important elements of learning. Because students with high motivation become eager to learn and can exert more effort in situations where they have difficulty learning (Kramer et al., 2024). Therefore, the finding that mathematics teaching with metacognitive CRIME increases students' motivation is important for the mathematics education literature. As a result of the observation, it was observed that the activities implemented with metacognitive CRIME also increased the motivation of intermediate-level students to participate in the lesson. Özsoy (2007) stated that problem-solving success affects motivation. He concluded that as problem-solving success increases, motivation also increases. In this context, the fact that the metacognitive CRIME technique is an approach based on problem-solving can be interpreted as increasing the motivation of students who are successful at a moderate level.

The study results suggest that the activities implemented with metacognitive CRIME do not attract the attention of students with moderate and low achievement levels. Many studies have also demonstrated this difference between high and low-level students. For example, Öztürk (2021) revealed that metacognitive training has different effects on high-achieving and low-achieving students in his research using the IMPROVE technique. Özdoğan (2024) also compared the metacognitive skills of good and poor problem solvers to understand the problem, revealing that the metacognitive skills of good and poor problem solvers differed significantly. In this respect, it can be said that the study supports the literature. The differentiation of metacognitive skills in good and poor students may lead to thoughts such as matching this skill with the characteristics of successful or good students (Chen & McDunn, 2022; Eberhart et al., 2024; Jeong & Kim, 2025). However, metacognitive skills can be developed in students with different cognitive domain levels.

The study results showed that students with high and medium levels of success thought that their problem-solving skills improved during the learning process with the metacognitive CRIME technique. According to the information obtained from the

meta-analysis study, Aydın and Öztürk (2024) concluded that the mathematical achievement of students who received metacognitive training increased, and their problem-solving success was significantly higher. In addition, many studies in the literature have shown that the problem-solving success of students who received metacognitive training increased (Jeong & Kim, 2025; Özçakır -Sümen, 2021). In addition, Öztürk, Sarıkaya, and Ada-Yıldız (2024) determined as a result of the research that metacognition is an important predictor of problem-solving. In this context, it can be said that the finding that the applications carried out with the metacognitive CRIME technique improved students' problem-solving skills supports the literature. The interviews determined that students with low success thought there was no development in their problem-solving skills. The difference between high and low-success students may have arisen from the nature of metacognition. Öztürk et al. (2024) pointed out that even if students have sufficient knowledge and skills, they will not be successful in problem-solving if their metacognitive skills are not sufficiently developed. Therefore, it can be thought that low-achieving students use negative expressions about their problem-solving skills because their metacognitive skills are not sufficiently developed.

The researcher's observations determined that students who learned with the metacognitive CRIME technique were more systematic in solving problems and that students who had difficulty determining strategies were more successful. Tabuyo (2024) stated that students with advanced metacognitive skills can be more successful in determining strategies in the problem-solving process. Since metacognition inherently requires the individual's planning, monitoring, and evaluation skills, individuals with advanced metacognitive skills are expected to become systematic in problem-solving.

As a result of the interviews, it was determined that students with high and medium levels of success thought that their course success increased during the learning process with the metacognitive CRIME technique. Baten et al. (2017) stated that metacognitive education was more effective in mathematics learning, even in young children. Aydın and Öztürk (2024) concluded in their meta-analysis that students who received metacognitive training were more successful in mathematics lessons. Eberhart et al. (2024) stated that the self-efficacy of students who received metacognitive education increased. In this context, it can be said that the results obtained support the literature. Many studies have also shown that metacognitive training increases course success (Öztürk, 2021; Şahinkaya et al., 2022). In this context, it can be said that the student's thoughts that their course success increased to support the literature.

Studies have found that students with low and medium levels of success think there is no improvement in their course success. This difference between high and low-level students has also been revealed in many studies. For example, Mert and Baş (2019) stated that as students' metacognitive awareness decreases, their mathematics success also decreases. Therefore, it can be thought that students with medium and low levels of success use negative expressions about their course success because their metacognitive skills are not sufficiently developed.

The researcher's observations indicated that students who learned with the metacognitive CRIME technique participated in the lesson with better equipment, made more effort to learn, and made the learning process more efficient. Since metacognition requires the individual's planning, monitoring, and evaluation skills, students who have developed these skills are expected to become systematic in problem-solving. As a result, their success in the lesson will increase.

As a result of the interviews, it was determined that students with high levels of success developed their monitoring skills in the learning process with the metacognitive CRIME technique. In contrast, intermediate-level students thought these skills were partially developed. Okumuş and Öztürk (2024) stated that students gained skills in cognitive strategy creation, cognitive foresight, cognitive foresight monitoring, cognitive awareness, and motivation in the teaching environment designed based on metacognitive planning and monitoring. Başokçu and Güzel (2021) concluded that metacognitive monitoring skills were important in evaluating the PISA mathematics test. In this context, it can be said that the literature supports the findings that the applications made with the metacognitive CRIME technique developed students' metacognitive monitoring skills in the learning process.

The researcher stated in his observations that the metacognitive CRIME technique enabled students to become aware of the information they had while allowing them to determine the issues they were lacking and to make inferences accordingly. As a result of the interviews, it was determined that the participants who showed high levels of success developed their metacognitive assessment skills during the learning process with the metacognitive CRIME technique. At the same time, the intermediate-level students thought they had partially developed. While the successful students overlooked their mistakes and deficiencies before the application, they started questioning and checking each step after the application. The students generally understood how they could learn, but some had difficulty evaluating the process. This situation shows that metacognitive awareness develops over time and progresses differently in each student.

CONCLUSION AND RECOMMENDATIONS

Middle school students with the metacognitive CRIME technique on ratio-proportion and percentages generally support the literature but provide some original results. The findings in the study that the metacognitive CRIME technique increases motivation, problem-solving, course success, and metacognitive skills support the literature. The original result of the study is that the metacognitive CRIME technique improves metacognitive monitoring and evaluation skills in high-achieving students. At the same time, it does not show any change in the metacognitive skills of medium and low-achieving students. In addition, the fact that the metacognitive CRIME technique is effective in the problem-solving skills of high -and medium-achieving students, while it does not show any change in the problem-solving skills of low-achieving students, is another original result of the study.

Implications and Future Research

This study was conducted to examine how the metacognitive CRIME technique affects the teaching of the subject of ratio-proportion and percentages and was conducted under some limitations. The first of these limitations is related to the study model. The study was conducted according to the case study model, and data were collected only through observation and interviews. In this study, participant evaluations were prioritized, and documents were not included due to the nature of the research questions. Including documents by future researchers may provide diversification of data collection tools. This study was designed using qualitative research; only generalizations were made about the participants. Future researchers can also use quantitative data tools to conduct studies in mixed research designs and make generalizations.

The results obtained in the study show that although the metacognitive CRIME technique improved the metacognitive skills of students with high success in teaching ratio-proportion and percentages, it did not show the targeted effect on the metacognitive skills of students with medium and low success. Therefore, it can be suggested that future researchers should conduct a study to develop a teaching technique to improve students' metacognitive skills with medium and low success. This research is expected to include an application process and contribute to mathematics teachers and mathematics education researchers in terms of applying and evaluating the metacognitive CRIME technique.

Declaration of Conflicting Interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

This study was supported by Bayburt University Scientific Research Project Coordinatorship under Grant Number 2024/69001-01-01. The authors thank Bayburt University Scientific Research Project Coordinatorship for their support.

Statements of publication ethics

We hereby declare that the study has not unethical issues and that research and publication ethics have been observed carefully.

Researchers' contribution rate

The study was conducted and reported with equal collaboration of the researchers.

Ethics Committee Approval Information

All legal permits were obtained for the study. The study reported in this paper was conducted under the approval of the Ethics Committee of Bayburt University (E-83541712-050.99-138817). Informed consent was obtained from the participants included in the study. The study was conducted with voluntary students.

REFERENCES

- Ares Ferreirós, M., Alfonso Gil, S., Rodríguez Enríquez, M., Conde Rodríguez, Á., & Deaño Deaño, M. (2024). Improvement in simultaneous processing through metacognitive instruction. *Frontiers in Education*, 9, (p. 1346739).
- Aydın, Ü., & Öztürk, M. (2024). Üstbilişe dayalı öğretimin matematik başarısı ve problem çözmeye etkisi: Bir meta-analiz araştırması. *Manisa Celal Bayar University Journal of The Faculty of Education*, 12(2), 286–304. <https://doi.org/10.52826/mcbuefd.1450124>
- Baumanns, L., & Rott, B. (2022). The process of problem posing: Development of a descriptive phase model of problem posing. *Educational Studies in Mathematics*, 110(2), 251-269. <https://doi.org/10.1007/s10649-021-10091-0>
- Baten, E., Praet, M., & Desoete, A. (2017). The relevance and efficacy of metacognition for instructional design in the domain of mathematics. *ZDM Mathematics Education*, 49, 613–623. <https://doi.org/10.1007/s11858-017-0851-y>
- Chen, S., & McDunn, B. A. (2022). Metacognition: History, measurements, and the role in early childhood development and education. *Learning and Motivation*, 78, 101786. <https://doi.org/10.1016/j.lmot.2022.101786>
- Eberhart, J., Ingendahl, F., & Bryce, D. (2025). Are metacognition interventions in young children effective? Evidence from a series of meta-analyses. *Metacognition and Learning*, 20(7), 1-45. <https://doi.org/10.1007/s11409-024-09405-x>
- Erdem, A., & Öztürk, M. (2023). Manipülatif destekli üstbilişsel planlamaya dayalı öğrenme ortamı tasarımı: Çarpanlar ve katları konusu örneği. *Bolu Abant İzzet Baysal Üniversitesi Eğitim Fakültesi Dergisi*, 23(2), 559-584. <https://doi.org/10.17240/aibuefd.2023.-1099309>
- Flavell, J. H. (1979). Metacognition and cognitive monitoring: A new area of cognitive-developmental inquiry. *American Psychologist*, 34(10), 906-911. <https://doi.org/10.1037/0003-066X.34.10.906>
- Hythecker, V. I., Dansereau, D. F., & Rocklin, T. R. (1988). An analysis of the processes influencing the structured dyadic learning environment. *Educational Psychologist*, 23(1), 23-37.
- Jeong, J., & Kim, S. (2025). Young children's metacognition in problem-solving through question-asking. *Journal of Experimental Child Psychology*, 254, 106207. <https://doi.org/10.1016/j.jecp.2025.106207>
- Gandolfi, E., Limata, T., Favatà, R., & Iani, F. (2024). Does peer tutoring have negative effects? An investigation and intervention on tutors' implicit theories and beliefs of intelligence. *Educational Psychology*, 44(5), 632-648.
- Kramer, A. W., Huizenga, H. M., Van Duijvenvoorde, A. C., & Krabbendam, L. (2024). Do I want to learn today? Day-to-day variations in adolescents' academic motivation and effort. *Learning and Motivation*, 85, 101957. <https://doi.org/10.1016/j.lmot.2023.101957>
- Kwang, T. S. (2000). *The effect of metacognitive training on the mathematical word problem solving of Singapore 11–12 year olds in a computer environment* (Doctoral dissertation). University of Leeds.
- McCormick, S., & Cooper, J. O. (1991). Can SQ3R facilitate secondary learning disabled students' literal comprehension of expository tests? Three experiments. *Reading Psychology*, 12, 239–271.

- Merriam, S. B. (2013). *Nitel araştırma: Desen ve uygulama için bir rehber* (3. baskıdan çeviri). (Çeviri Editörü: S. Turan). Ankara: Nobel Yayın Dağıtım.
- Mert, M., & Baş, F. (2019). Ortaokul öğrencilerinin matematiğe yönelik kaygı, üstbilişsel farkındalık düzeyleri ve ilgili değişkenlerin matematik başarılarındaki etkisi. *Türk Bilgisayar ve Matematik Eğitimi Dergisi*, 10(3), 732-756. <https://dx.doi.org/10.16949/turkbilmat.508347>
- Mevarech, Z. R., & Kramarski, B. (1997). IMPROVE: A multidimensional method for teaching mathematics in heterogeneous classrooms. *American Educational Research Journal*, 34(2), 365-395. <https://doi.org/10.3102/00028312034002365>
- Okumuş, Ö., & Öztürk, M. (2024). Metacognitive training for algebra teaching to high school students: An action research study. *Learning and Motivation*, 88, 102064. <https://doi.org/10.1016/j.lmot.2024.102064>
- Özdoğan, E. G. (2024). *Metacognitive planning skills of good and poor problem solvers towards understanding the problem*. (Publication No. 928139) [Master Thesis, Bayburt University-Bayburt]. Council of Higher Education Thesis Center.
- Özçakır Sümen, Ö. (2021). The mediating role of metacognitive self-regulation skills in the relationship between problem-posing skills and mathematics achievement of primary pre-service teachers. *International Online Journal of Education and Teaching (IOJET)*, 8(3), 2081-2096. <https://doi.org/10.15345/iojet.2021.03.129>
- Öztürk, M. (2021). An embedded mixed method study on teaching algebraic expressions using metacognition-based training. *Thinking Skills and Creativity*, 41, 100787. <https://doi.org/10.1016/j.tsc.2021.100787>
- Öztürk, M., Sarıkaya, İ., & Ada Yıldız, K. (2024). Middle school students' problem solving performance: Identifying the factors that influence it. *International Journal of Science and Mathematics Education*, 22(6), 1363-1379. <https://doi.org/10.1007/s10763-023-10423-5>
- Prins, F. J., Veenman, M. V., & Elshout, J. J. (2006). The impact of intellectual ability and metacognition on learning: New support for the threshold of problematicity theory. *Learning and instruction*, 16(4), 374-387. <https://doi.org/10.1016/j.learninstruc.2006.07.008>
- Şahinkaya, T., Öztürk, M., & Albayrak, M. (2022). Üstbilişsel IMPROVE tekniğinin oran-orantının öğretimi ve orantısız akıl yürütme becerisinin geliştirilmesi üzerine etkisi. *Kocaeli Üniversitesi Eğitim Dergisi*, 5(2), 495-516. <https://doi.org/10.33400/kuje.1137016>
- Tabuyo, A. T. (2024). Metacognition and mathematical problem-solving performance of pre-service teachers. *Pegem Journal of Education and Instruction*, 14(4), 474-480. <https://doi.org/10.14527/pegegog.2024.005>
- Teong, S. K. (2003). The effect of metacognitive training on mathematical word-problem solving. *Journal of Computer Assisted*, 19(1), 46-55. <https://doi.org/10.1046/j.0266-4909.2003.00005.x>



| Research Article / Araştırma Makalesi |

Investigation of Cross-Cultural Measurement Invariance through Multi-Group Confirmatory Factor Analysis and Multi-Group Latent Class Analysis

Kültürlerarası Ölçme Değişmezliğinin Çok Gruplu Doğrulatoryı Faktör Analizi ve Çok Gruplu Örtük Sınıf Analiziyle İncelenmesi¹

Fatma COŞKUN², Hamide Deniz GÜLLEROĞLU³

Keywords

1. Measurement invariance
2. Confirmatory factor analysis
3. Latent class analysis
4. TIMSS 2019
5. Student surveys

Anahtar Kelimeler

1. Ölçme değişmezliği
2. Doğrulatoryı faktör analizi
3. Örtük sınıf analizi
4. TIMSS 2019
5. Öğrenci anketleri

Received/Başvuru Tarihi
04.01.2025

Accepted / Kabul Tarihi
10.10.2025

Abstract

Purpose: This study aims to examine cross-cultural measurement invariance using multi-group confirmatory factor analysis (MG-CFA) and multi-group latent class analysis (MG-LCA).

Design/Methodology/Approach: For this purpose, data obtained from the 'Students Like Learning Mathematics Scale' in the TIMSS 2019 study were used. The sample of the research was determined using the maximum variation sampling method. Measurement invariance analyses were conducted on 15 comparison groups formed by Singapore, Hong Kong, Japan, Norway, Turkey, and South Africa.

Findings: The MG-CFA results showed that strict measurement invariance was achieved only between Singapore-Hong Kong and Hong Kong-Norway. Between South Africa and Turkey, measurement invariance was observed at the level of structural invariance, while in other groups, it was achieved at the level of metric invariance. According to the MG-LCA results, measurement invariance was established at the partially homogeneous model level for some groups and at the heterogeneous model level for others.

Highlights: In MG-LCA, it was determined that for the cross-cultural comparisons where measurement invariance was achieved at the partially homogeneous model level, measurement invariance was largely achieved at the metric invariance level according to MG-CFA. The MG-CFA and MG-LCA approaches address measurement invariance from different perspectives. MG-CFA is a sample- and data-oriented approach, whereas MG-LCA is a person-centered approach. Using these two methods together provides more comprehensive insights into why measurement invariance could not be achieved.

Öz

Çalışmanın amacı: Bu çalışmada, kültürlerarası ölçme değişmezliğinin çok gruplu doğrulatoryı faktör analizi (ÇG-DFA) ve çok gruplu örtük sınıf analizi (ÇG-ÖSA) ile incelenmesi amaçlanmıştır.

Materyal ve Yöntem: Bu amaç doğrultusunda, TIMSS 2019 uygulamasında "Matematik Öğrenmeyi Sevmek Ölçeği" ile elde edilen veriler kullanılmıştır. Araştırmanın örneklemini, maksimum çeşitlilik yöntemi ile belirlenmiştir. Ölçme değişmezliği analizleri Singapur, Hong Kong, Japonya, Norveç, Türkiye ve Güney Afrika Cumhuriyeti'nin oluşturduğu 15 karşılaştırma grubu üzerinde gerçekleştirilmiştir.

Bulgular: ÇG-DFA sonuçları, katı ölçme değişmezliğinin yalnızca Singapur-Hong Kong ve Hong Kong-Norveç arasında sağlandığını göstermiştir. Güney Afrika Cumhuriyeti ve Türkiye arasında ölçme değişmezliği yapısal değişmezlik düzeyinde kalırken; diğer gruplarda arasında ölçme değişmezliğinin metrik değişmezlik düzeyinde sağlandığı görülmüştür. ÇG-ÖSA sonuçlarına göre, bazı gruplar arasında kısmi homojen model, diğerleri arasında ise heterojen model düzeyinde ölçme değişmezliği sağlanmıştır.

Önemli Vurgular: ÇG-ÖSA' da ölçme değişmezliğinin kısmi homojen model düzeyinde sağlandığı kültürlerarasında ÇG-DFA'ya göre ölçme değişmezliğinin büyük oranda metrik değişmezlik düzeyinde sağlandığı belirlenmiştir. ÇG-DFA ve ÇG-ÖSA yaklaşımları ölçme değişmezliğini farklı boyutlardan ele almaktadır. ÇG-DFA, örneklem ve veri odaklı bir yaklaşımdır; ÇG-ÖSA birey merkezli bir yaklaşımdır. Bu iki yöntemin birlikte kullanılması, ölçme değişmezliğinin neden sağlanamadığı hakkında daha kapsamlı bilgiler sağlamaktadır.

¹ This study is based on the doctoral dissertation of the first author under the supervision of the second author.

² Corresponded Author, Kahramanmaraş Sütçü İmam University, Faculty of Education, Department of Educational Measurement and Evaluation, Kahramanmaraş, Türkiye; <https://orcid.org/0000-0002-6388-3504>

³ Ankara University, Ankara Faculty of Education, Department of Educational Measurement and Evaluation, Ankara, Türkiye; <https://orcid.org/0000-0001-7919-0596>

INTRODUCTION

In recent years, international large-scale assessments have gained increasing importance in educational research. Until the late 1950s, systematic data on educational outcomes were seldom collected at either the national or international level. However, the past decade has seen significant growth in this field (Kirsch & Braun, 2020). Notably, the Trends in International Mathematics and Science Study (TIMSS) and the Progress in International Reading Literacy Study (PIRLS), conducted by the International Association for the Evaluation of Educational Achievement (IEA), as well as the Programme for International Student Assessment (PISA), carried out by the Organisation for Economic Co-operation and Development (OECD), are at the forefront of large-scale assessment activities (Hernandez-Torrano & Courtney, 2021). These large-scale assessments capture the attention of the media and the public and play a pivotal role in shaping education policies worldwide (Johansson, 2016). Additionally, many governments make significant financial investments to realign their education policies based on the results of these large-scale assessments (Hambleton & Zenisky, 2011).

Large-scale assessments typically require cross-border measurements conducted through multiple testing tools (Erikcan et al., 2023). In general, large-scale assessments contribute to the development of cross-cultural understanding by comparing student characteristics across different countries (Davidov et al., 2014). One of the primary objectives of large-scale assessments, however, is to compare student characteristics at an international level. This inherently necessitates comparisons across multiple cultures (Hernandez-Torrano & Courtney, 2021). Such comparisons can only be achieved by adapting the tests developed within the scope of large-scale assessments to different languages and cultures (Hambleton & Zenisky, 2011). Tests adapted to various languages and cultures as part of large-scale assessments bring challenges in terms of cross-cultural validity and comparability, such as construct validity, test equivalence, and equivalence of testing conditions (Erikcan et al., 2023). Even when translators achieve absolute accuracy, the cultural context of words used in the source and target languages may lead to differing interpretations (Hambleton & Zenisky, 2011). The lack of equivalence between test versions adapted for different cultures may stem from a variety of factors, including geographic features, traditions, societal moral values, levels of economic development, religious beliefs, or education systems (Hambleton, 2004). To ensure comparable measurement results across different cultures, cultural biases must be minimized (Erikcan et al., 2023). This is achievable through the establishment of measurement invariance. Ignoring measurement invariance significantly jeopardizes the validity of comparison-based analyses (Vandenberg & Lance, 2000).

Failure to properly test the measurement invariance of data obtained from different cultures or nations can lead to the misinterpretation of measurement results (Meitinger et al., 2020). Therefore, before making significant inferences based on cross-cultural data, it is essential to determine whether the relevant constructs are measured in the same way across different cultures (Cieciuch et al., 2019). Van de Vijver (1998) compared comparisons without establishing measurement invariance to comparing apples and oranges, arguing that such comparisons are meaningless and incorrect. Direct comparisons made using terms that may vary across countries, such as the concept of "middle class" commonly used in educational sciences, can result in erroneous conclusions. For instance, the standards of the "middle class" in a country with a high socio-economic status may differ significantly from those in a country with a lower socio-economic status. As a result, studies that do not provide evidence of measurement invariance carry a risk of misinterpreting analysis results (Van de Vijver & Leung, 2021). In this context, education researchers must carefully examine measurement invariance when conducting studies based on international comparisons. For example, Wendt et al. (2017), in a study testing four different levels of measurement invariance using data from 37 countries in the TIMSS and PIRLS applications, found that the model allowing for country-specific measurement structures provided the best fit to the data. This finding highlights the necessity of ensuring measurement invariance for the accuracy of comparative analyses in large-scale applications (Borsboom, 2006).

As understood from the previous explanations, measurement invariance fundamentally refers to the ability of a measurement tool to represent the same construct across all subgroups to which it is applied (Davidov et al., 2014). The literature includes numerous definitions and varied terminologies related to measurement invariance. In his comprehensive study, Johnson (1998) identified more than 50 definitions of "measurement invariance." Additionally, it is observed that the term "measurement equivalence" is also conceptually used in place of measurement invariance in the literature (Kankaras, 2010). Despite the abundance of definitions and varying terminologies in the literature, the common focus of measurement invariance studies lies in the feasibility of making comparisons across groups (Vandenberg & Lance, 2000).

Statistically, measurement invariance is tested within a hierarchical structure, progressing from a model where no parameters are constrained to a model where all parameters are constrained, based on measurements obtained from two different groups (Meredith, 1993; Thissen, 2001). Jöreskog (1971) first investigated statistical evidence for measurement invariance using Multi-Group Confirmatory Factor Analysis (MG-CFA). Subsequently, Sörbom's adaptations of the LISREL program for MG-CFA in 1974 and 1978 contributed to the widespread use of the MG-CFA technique and guided its inclusion in many of today's analysis programs (Van De Schoot et al., 2015). Although statistical analyses of measurement invariance appear to have originated with MG-CFA, many new techniques have been developed over time. Today, numerous methods are employed to test measurement invariance, including MG-CFA, Multi-Group Latent Class Analysis (MG-LCA), Multilevel Confirmatory Factor Analysis, Multilevel Factor Mixture Modeling, Bayesian approaches, Alignment Optimization, and Item Response Theory (IRT)-based models (Kim et al., 2017; Lubke & Neale, 2008). However, the literature shows that MG-CFA, MG-LCA, and IRT-based techniques are the most commonly used methods in measurement invariance studies (Davidov et al., 2014). MG-LCA, in particular, is considered an

alternative to MG-CFA by many researchers due to its flexible assumptions (Güngör et al., 2013; Kankaras et al., 2011). MG-LCA was introduced approximately 14 years after the initial use of MG-CFA, in Clogg and Goodman's 1985 paper titled "Simultaneous Latent Structure Analysis in Several Groups." Developed to evaluate measurement invariance in data obtained from different groups, MG-LCA shares many similarities with MG-CFA (Clogg & Goodman, 1985).

In MG-LCA, as in MG-CFA, there is a hierarchical structure progressing from a model where no parameters are constrained across groups to a model where all parameters are constrained. In MG-LCA, the model where all parameters are freely estimated among the compared groups is referred to as the heterogeneous model, the model where certain parameters are fixed is called the partially homogeneous model, and the model where all parameters are fixed is known as the homogeneous model (Clogg & Goodman, 1985; Kankaras & Vermunt, 2014). Despite sharing many theoretical and practical similarities, MG-CFA and MG-LCA also differ in several aspects (Kankaras et al., 2011). For instance, although both techniques are based on latent variable modeling, MG-CFA assumes that latent and observed variables are continuous, whereas MG-LCA treats both latent and observed variables as categorical (Collins & Lanza, 2009; Kankaras et al., 2011). Consequently, MG-CFA requires the fulfillment of univariate and multivariate normality assumptions (Brown & Moore, 2012), while these assumptions are not examined in MG-LCA (Kankaras et al., 2011). MG-LCA operates with more flexible assumptions, with its primary assumption being the principle of local independence (Eid et al., 2003).

From another perspective, while MG-CFA represents a variable-centered approach, MG-LCA is categorized among person-centered approaches (Bergman & Magnusson, 1997; Bergman & Wangby, 2014). This distinction arises from the analytical approaches of the two techniques. In MG-CFA, the focus is on evaluating whether factor structures are similar across groups and identifying potential differences between groups. In contrast, MG-LCA assigns each participant to a specific class based on their response patterns, and the analysis is conducted by examining the differences and similarities among these participant-formed classes (Brown & Moore, 2012; Kankaras et al., 2011). In other words, while MG-CFA involves building a model for a latent trait, MG-LCA identifies latent classes based on participants' response patterns.

When examining measurement invariance studies, it is observed that MG-CFA is the most commonly used technique (Yandi et al., 2015). On the other hand, Eid et al. (2003) have presented compelling arguments regarding the advantages of using MG-LCA to investigate measurement invariance in cross-cultural comparisons. Indeed, numerous subsequent studies (Janousch et al., 2022; Kankaras et al., 2018; Zhao & Jin, 2023) have demonstrated the effectiveness of MG-LCA in cross-cultural measurement invariance studies. Building on this context, this study aims to examine cross-cultural measurement invariance using MG-CFA and MG-LCA techniques with data from TIMSS 2019. A review of the literature reveals that studies employing both techniques together in cross-cultural measurement invariance research are quite limited. From this perspective, this study is expected to contribute to the field. Furthermore, conducting cross-cultural measurement invariance research using TIMSS 2019 data is of significant importance for highlighting the relationship between educational research and cross-cultural measurement invariance. In line with this overarching goal, this study seeks to answer the following questions.

1. Does the Students Like Learning Mathematics Scale ensure measurement invariance across Singapore, Hong Kong, Japan, Norway, Turkey, and South Africa according to the MG-CFA method?
2. Does the Students Like Learning Mathematics Scale ensure measurement invariance across Singapore, Hong Kong, Japan, Norway, Turkey, and South Africa according to the MG-LCA method?

METHOD/MATERIALS

Research Design

This study aims to examine cross-cultural measurement invariance using the Students Like Learning Mathematics Scale (SLLMS) from the TIMSS 2019 assessment, employing MG-CFA and MG-LCA techniques. In this respect, it is characterized as a descriptive study since it seeks to examine the current situation without altering it (Fraenkel & Wallen, 2009).

Study Group

In this study, data from the Students Like Learn Mathematics Scale administered to 4th-grade students in the TIMSS 2019 assessment were used to apply MG-CFA and MG-LCA techniques. Accordingly, the population of this study consists of the 58 countries that participated in the 4th-grade level of TIMSS 2019 (Mullis et al., 2020). However, conducting analyses for all 58 countries in a single study was not feasible due to time and scope limitations. Therefore, countries representing high, medium, and low achievement levels, and differing significantly both geographically and culturally, were selected from among the participating nations. To achieve this, the maximum variation sampling method, a subtype of purposive sampling, was employed, and the analyses were conducted using data from Hong Kong, South Africa, Japan, Norway, Singapore, and Turkey (Gliner et al., 2017). Table 1 presents the student proportions from the countries included in the study sample.

Table 1. Student Proportions of the Countries in the Study Sample

Country	Frequency (f)	Percentage (%)
Hong Kong	3386	9
South Africa	11891	34
Japan	4196	12

Country	Frequency (f)	Percentage (%)
Norway	4526	13
Turkey	4599	13
Singapore	6839	19
Total	35437	100

Data Collection Tools

Students Like Learn Mathematics Scale

SLLMS, included in the TIMSS 2019 assessment, consists of nine items designed to measure students' attitudes toward learning mathematics. The scale is structured as a four-point Likert scale, with items rated as "agree a lot," "agree a little," "disagree a little," and "disagree a lot." Two of the items on the scale are negatively worded and require reverse scoring. This scale was first introduced in the TIMSS assessment in 2011 (Yin & Fishbein, 2020). The data for the SLLMS used in the MG-CFA and MG-LCA analyses in this study were accessed through the IEA's website (<https://timssandpirls.bc.edu/timss2019/>). The confirmatory factor analysis indices for the SLLMS data used in the study are presented in Table 2.

Table 2. Indices Related to the CFA Results of the SLLMS

Country	χ^2	sd	χ^2/sd	RMSEA	CFI	GFI	SRMR	NFI	IFI
Singapore	1509.09	27	55.89	.09	.96	.95	.03	.96	.96
Hong Kong	1337.89	27	49.55	.12	.94	.92	.05	.94	.94
Japan	1322.54	27	48.98	.11	.95	.93	.05	.95	.95
Norway	874.40	27	32.38	.09	.97	.95	.03	.97	.97
Turkey	1101.28	27	40.78	.09	.94	.95	.05	.94	.94
South Africa	2473.93	27	91.62	.09	.91	.95	.06	.91	.91

Table 2 shows that the fit indices obtained from the CFA of the SLLMS (CFI > .90, GFI > .90, NFI > .90, IFI > .90) indicate an acceptable model fit for the six countries in the study sample (Kline, 2016).

Data Analysis

This study aimed to examine cross-cultural measurement invariance using the data obtained from the SLLMS through MG-CFA and MG-LCA techniques. Before proceeding to the main analyses, the data from Hong Kong, South Africa, Japan, Norway, Singapore, and Turkey were compiled into a single SPSS dataset and examined for missing data and outliers. It was determined that the rate of missing data ranged from 0% to 100% for participants and from 1.6% to 4.4% for scale items. Consequently, 273 participants who left all items unanswered and 1,200 participants with more than 30% missing responses were excluded from the dataset. As recommended by Kline (2016) and Tabachnick and Fidell (2019), the expectation-maximization algorithm was used to address missing data. Univariate outliers were detected based on z-scores, whereas multivariate outliers were identified using Mahalanobis distance (Tabachnick & Fidell, 2019). Following these analyses, data from 33,357 participants were included in the analysis. In tests of normality and multicollinearity assumptions, it was found that the normality assumption was met through skewness and kurtosis coefficients as well as graphical examinations (Demir, 2019). Additionally, multicollinearity assessments indicated that VIF, CI, and tolerance values (VIF < 10; CI < 30; Tolerance > 10) were within acceptable limits (Tabachnick & Fidell, 2019). Exploratory factor analysis was used to assess unidimensionality and local independence. It was determined that the SLLMS exhibited a unidimensional structure, thereby satisfying the assumption of local independence (Demars, 2016). IBM SPSS Statistics (Version 20.0) was used for data preparation. AMOS (Version 24.0) was employed for MG-CFA, and Latent GOLD (Version 6.0) was used for MG-LCA.

The MG-CFA and MG-LCA techniques used to determine cross-cultural measurement invariance were applied to all pairwise comparisons among the six countries comprising the study sample. This approach helps identify varying levels of measurement invariance that may exist among subgroups when full measurement invariance cannot be established across all cultures. Furthermore, it provides more detailed information for cross-cultural comparisons (Davidov et al., 2014). Similarly, Wu et al. (2007), in their study examining cross-cultural measurement invariance using TIMSS data, investigated measurement invariance across 21 comparison groups representing all pairwise comparisons among the seven countries they identified.

In MG-CFA, four levels of measurement invariance were tested, progressing from a model with no parameter constraints to a model with all parameters constrained (Wu et al., 2007). To determine the best-fitting measurement invariance model, several fit indices were considered. The first of these is the " $\Delta\chi^2$ " value, based on the difference in chi-square values between two models. While the chi-square difference test ($\Delta\chi^2$) is still widely used, it has been found to be sensitive to sample size. Therefore, it is recommended not to rely solely on the $\Delta\chi^2$ index for decisions regarding measurement invariance (Cheung & Rensvold, 2002; Kline, 2016). In subsequent years, researchers developed the Comparative Fit Index (CFI) and the Root Mean Square Error of Approximation (RMSEA), which are less affected by sample size, for use in MG-CFA evaluations (Wu et al., 2007). In this study, the $\Delta\chi^2$, Δ CFI, and Δ RMSEA indices were used to determine the level of measurement invariance in model comparisons during MG-CFA analyses. Accordingly, in model evaluations progressing from structural measurement invariance to strict measurement invariance, at least two of the following criteria were required to be met: Δ RMSEA $\leq$.05, Δ CFI $\leq$ -.01, and the $\Delta\chi^2$ difference not being significant (Cheung & Rensvold, 2002; Wu et al., 2007).

In MG-LCA, to determine measurement invariance, the first step involved identifying the latent class model representing the six countries in the study sample (Magidson et al., 2020). Determining the number of classes in latent class models is based on principles of parsimony and interpretability, similar to exploratory factor analysis, alongside data reduction (Güngör et al., 2013). Although various indices exist for identifying latent class models, researchers primarily use the Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC). BIC is more suitable for large samples, whereas AIC is preferable for smaller samples. These indices identify the optimal model based on the number of classes analyzed: the model with the lowest AIC and BIC values as the number of classes increases is considered the best fit (Oberski, 2016). Latent class models, like MG-CFA, also use the chi-square index. However, since the chi-square index is sensitive to sample size in MG-LCA, the use of the BIC index is recommended as an alternative (Magidson et al., 2020). In this study, as the sample sizes analyzed were large enough to affect the chi-square values, the BIC index was used to evaluate the MG-LCA models.

FINDINGS

The findings of the study are presented in two sections. First, MG-CFA was conducted for all pairwise comparison groups among the cultures of Hong Kong, South Africa, Japan, Norway, Singapore, and Turkey, which constitute the study sample. Subsequently, MG-LCA was applied to all comparison groups analyzed in MG-CFA.

Findings Related to MG-CFA

Measurement invariance analyses were conducted across 15 groups based on pairwise comparisons of the countries constituting the sample. These groups are as follows: Singapore-Hong Kong, Singapore-Japan, Singapore-Norway, Singapore-Turkey, Singapore-South Africa, Hong Kong-Japan, Hong Kong-Norway, Hong Kong-Turkey, Hong Kong-South Africa, Japan-Norway, Japan-Turkey, Japan-South Africa, Norway-Turkey, Norway-South Africa, and Turkey-South Africa. Measurement invariance analyses in MG-CFA started by testing the structural invariance model, which imposes the fewest constraints on model parameters. Subsequently, for comparison groups that met the requirements of the next higher model, metric invariance, scalar invariance, and strict invariance models were tested in sequence to conclude the analyses. Table 3 presents the results for the structural invariance model, the initial level of measurement invariance tested in MG-CFA, across the 15 comparison groups.

Table 3. Indices Regarding the Structural Invariance Level of MG-CFA

Comparison Groups	N	χ^2	<i>p</i>	RMSEA	CFI
Singapore-Hong Kong	9847	1379.54	.00	.05	.98
Singapore-Japan	10741	1344.11	.00	.05	.98
Singapore-Norway	10812	1477.89	.00	.05	.98
Singapore-Turkey	11110	1300.86	.00	.05	.98
Singapore-South Africa	17359	1256.77	.00	.04	.98
Hong Kong-Japan	7332	1042.35	.00	.05	.98
Hong Kong-Norway	7403	1176.11	.00	.05	.98
Hong Kong-Turkey	7701	999.11	.00	.05	.98
Hong Kong-South Africa	13950	955.05	.00	.04	.98
Japan-Norway	8297	1440.70	.00	.05	.98
Japan-Turkey	8595	963.69	.00	.05	.98
Japan-South Africa	14844	919.61	.00	.03	.98
Norveç-Turkey	8666	1097.45	.00	.05	.98
Norveç-South Africa	14915	1053.39	.00	.04	.98
Turkey-South Africa	15213	876.35	.00	.03	.98

Table 3 presents the χ^2 , RMSEA, and CFI indices to determine whether the structural invariance model is achieved in measurement invariance based on MG-CFA. Upon examining the table, it is observed that the change in the χ^2 difference test is significant for all 15 comparison groups ($p < .05$). However, studies have shown that the χ^2 difference test is prone to significance in MG-CFA conducted with sample sizes larger than 6,000 (Meade et al., 2008), as also evidenced in other research (Cheung & Rensvold, 2002). Therefore, for this study, the CFI and RMSEA indices, which are unaffected by sample size, are more effective in evaluating the measurement invariance models in MG-CFA. Analyzing the information in Table 3, it is determined that the structural invariance model is met for all 15 comparison groups in the study based on the CFI and RMSEA indices. This indicates

that the scale structure represented by the SLLMS remains consistent across all comparison groups in terms of the number of subdimensions and the distribution of items to subdimensions (Kline, 2016).

When the structural invariance model is achieved among the groups examined in MG-CFA, analyses proceed to the next level, metric invariance (Wu et al., 2007). Since structural invariance was found to be met for all pairwise comparison groups, the metric invariance model was analyzed for all groups. The MG-CFA results for the metric invariance model are presented in Table 4.

Table 4. Indices Regarding the Metric Invariance Level of MG-CFA

Comparison Groups	ΔCFI	$\Delta RMSEA$	$\Delta \chi^2$	p
Singapore-Hong Kong	-0.01	0.00	408.38	.00
Singapore-Japan	-0.01	0.01	565.19	.00
Singapore-Norway	0.00	0.00	153.09	.00
Singapore-Turkey	0.00	0.00	236.51	.00
Singapore-South Africa	-0.01	0.01	675.94	.00
Hong Kong-Japan	0.00	0.00	132.04	.00
Hong Kong-Norway	-0.01	0.00	282.01	.00
Hong Kong-Turkey	-0.01	0.01	524.40	.00
Hong Kong-South Africa	0.00	0.00	100.85	.00
Japan-Norway	-0.01	0.01	442.94	.00
Japan-Turkey	-0.01	0.01	556.79	.00
Japan-South Africa	0.00	0.00	109.52	.00
Norway-Turkey	-0.01	0.00	203.75	.00
Norway-South Africa	-0.01	0.01	524.97	.00
Turkey-South Africa	-0.02	0.01	371.91	.00

Upon examining the information in Table 4, it is observed that the $\Delta \chi^2$ difference test indicates significant changes for all comparison groups ($p < .05$). However, due to the large sample sizes used in MG-CFA, the ΔCFI and $\Delta RMSEA$ indices provide more reliable results. Analyzing Table 4, it is evident that the metric invariance model is valid for all comparison groups except Turkey-South Africa based on the ΔCFI and $\Delta RMSEA$ indices ($\Delta CFI \leq 0.01$; $RMSEA \leq 0.015$). Accordingly, the MG-CFA results indicate that measurement invariance remains at the structural invariance level for the Turkey-South Africa comparison. Indices for the scalar invariance model are provided in Table 5 for the 14 comparison groups where metric invariance was achieved.

Table 5. Indices Regarding the Scalar Invariance Level of MG-CFA

Comparison Groups	ΔCFI	$\Delta RMSEA$	$\Delta \chi^2$	p
Singapore-Hong Kong	-0.01	0.00	408.513	.00
Singapore-Japan	-0.03	0.02	2051.57	.00
Singapore-Norway	-0.02	0.02	1661.08	.00
Singapore-Turkey	-0.04	0.02	2399.19	.00
Singapore-South Africa	-0.05	0.02	3413.36	.00
Hong Kong-Japan	-0.02	0.01	717.13	.00
Hong Kong-Norway	-0.01	0.01	614.16	.00
Hong Kong-Turkey	-0.04	0.02	1434.78	.00
Hong Kong-South Africa	-0.04	0.02	1743.71	.00
Japan-Norway	-0.03	0.02	1585.10	.00
Japan-Turkey	-0.05	0.03	2356.01	.00
Japan-South Africa	-0.08	0.04	4357.66	.00
Norway-Turkey	-0.04	0.02	1996.48	.00
Norway-South Africa	-0.06	0.03	3219.17	.00

Upon examining the information in Table 5, the $\Delta \chi^2$ difference test shows significant changes for all comparison groups ($p < .05$). According to the ΔCFI and $\Delta RMSEA$ indices, the scalar invariance model is valid between the cultures of Singapore-Hong Kong and Hong Kong-Norway ($\Delta CFI \leq 0.01$; $\Delta RMSEA \leq 0.015$). On the other hand, scalar invariance is achieved for Hong Kong-Japan only based on the $\Delta RMSEA$ index. From the information in Table 5, it can be concluded that scalar invariance is satisfied only for Singapore-Hong Kong and Hong Kong-Norway among the 14 comparison groups. Based on this, the next level, strict invariance, was analyzed for these two groups. For the remaining 12 comparison groups, measurement invariance was found to remain at the metric invariance level. The details of the strict invariance model applied to Singapore-Hong Kong and Hong Kong-Norway through MG-CFA are presented in Table 6.

Table 6. Indices Regarding the Scalar Invariance Level of MG-CFA

Comparison Groups	ΔCFI	$\Delta RMSEA$	$\Delta \chi^2$	p
Singapore-Hong Kong	-0.01	0.00	444.61	.00
Hong Kong-Norway	-0.01	0.00	415.31	.00

Upon examining the information in Table 6, the $\Delta\chi^2$ difference test indicates significant changes between the Singapore-Hong Kong and Hong Kong-Norway comparison groups ($p < .05$). According to the ΔCFI and $\Delta RMSEA$ indices, the strict invariance model is valid between the cultures of Singapore-Hong Kong and Hong Kong-Norway ($\Delta CFI \leq 0.01$; $\Delta RMSEA \leq 0.015$). Accordingly, it was determined that strict measurement invariance was achieved only between the cultures of Singapore-Hong Kong and Hong Kong-Norway among the 15 comparison groups analyzed using MG-CFA. Structural invariance was found to hold for Turkey-South Africa, while metric invariance was achieved for the remaining 12 comparison groups

Findings Related to MG-LCA

Measurement invariance analyses based on MG-LCA were conducted across 15 groups formed by all pairwise comparisons of the countries in the sample, similar to MG-CFA. To achieve this, it was first necessary to identify the latent class model that best fits the research data for MG-LCA. Accordingly, Latent Class Analysis (LCA) was initially performed to determine the most suitable latent class model for the research data. After identifying the most appropriate latent class model via LCA, the classes within the latent model were named based on their theoretical foundations and the constructs they represent. The process of determining the best-fitting model begins with a single-class latent model and continues by increasing the number of latent classes as long as degrees of freedom allow (Magidson & Vermunt, 2004; Magidson et al., 2020). The model with the lowest information criterion value, used as an evaluation metric, statistically represents the model with the best fit to the data. However, since LCA, like Exploratory Factor Analysis (EFA), is a data reduction method, other criteria such as the number of parameters and the interpretability of the model must also be considered when determining the appropriate model (Arıcıgil-Çılan, 2015; Green, 1952; Güngör-Culha, 2012). Accordingly, LCA was repeated for the six countries in the sample and for the entire dataset to identify the latent class model that best fits the data. The estimated model parameters for models ranging from a single-class model to a six-class model for data obtained through the Students Like Learn Mathematics Scale are presented in Table 7.

Table 7. Model Parameters Related to Latent Class Analysis

Model	Par. ¹	BIC(LL)	AIC(LL)	AIC3(LL)	L ²	sd	Class. ²
1. Class	27	708239.44	708012.23	708039.23	236485.24	33330	0.00
2. Class	55	588045.86	587583.04	587638.04	116000.04	33302	0.02
3. Class	83	558242.48	557544.04	557627.04	85905.05	33274	0.04
4. Class	111	545867.90	544933.84	545040.84	73238.85	33246	0.06
5. Class	139	539332.54	538162.86	538301.86	66411.86	33218	0.10
6. Class	167	534632.14	533226.83	533393.83	61419.84	33190	0.11

*Note: ¹ number of parameters, ² classification errors

According to Table 7, as the models progress from a single-class model to a six-class model, there is a noticeable decrease in BIC, AIC, and AIC3 values. This indicates that the data is not limited to a single-class model but instead points to the presence of multiple latent classes. This finding suggests that the data is suitable for latent class analysis (Magidson & Vermunt, 2004). In determining the most appropriate number of classes in the latent model, attention must be given not only to BIC, AIC, and AIC3 information criteria but also to the number of parameters and classification error (Altıntaş, 2016; Lin, 2006). While the decrease in information criteria is pronounced from the single-class model to the three-class model, it becomes much smaller when moving from the three-class model to the six-class model. This pattern can be visualized in a graph with the number of classes on the horizontal axis and information criteria on the vertical axis, similar to the scree plot used in exploratory factor analysis to determine the number of factors (Coşkun, 2023; Coşkun & Gülleroğlu, 2023; Çüm et al., 2020). Considering all this information, the three-class latent model, where the slope of the decrease in information criteria is the lowest, is considered the most reasonable model for explaining the data in terms of the number of parameters and classification error. The three-class latent model, which has a classification error of 4%, indicates that the classification process is largely accurate. Accordingly, the model provides evidence of a 96% likelihood of correct classification (Altıntaş, 2016). The conditional probability parameters of the three-class latent model are presented in Table 8.

Table 8. Conditional Probability Parameters for the Three-Class Latent Model

Item Code	Item Levels	1. Class (0.50)	2. Class (0.33)	3. Class (0.17)
ASBM02A	Strongly Agree	0.83	0.15	0.02
	A little Agree	0.06	0.78	0.16
	A little Disagree	0.02	0.16	0.83
	Strongly Disagree	0.07	0.05	0.88
ASBM02B	Strongly Agree	0.71	0.23	0.06
	A little Agree	0.17	0.58	0.25
	A little Disagree	0.24	0.43	0.33
	Strongly Disagree	0.49	0.17	0.34

	Strongly Agree	0.79	0.17	0.04
ASBM02C	A little Agree	0.19	0.62	0.19
	A little Disagree	0.15	0.51	0.34
	Strongly Disagree	0.40	0.14	0.46
	Strongly Agree	0.75	0.21	0.04
ASBM02D	A little Agree	0.18	0.61	0.21
	A little Disagree	0.09	0.29	0.62
	Strongly Disagree	0.26	0.16	0.58
	Strongly Agree	0.87	0.12	0.01
ASBM02E	A little Agree	0.05	0.87	0.08
	A little Disagree	0.01	0.29	0.70
	Strongly Disagree	0.04	0.04	0.92
	Strongly Agree	0.86	0.12	0.03
ASBM02F	A little Agree	0.33	0.59	0.08
	A little Disagree	0.07	0.49	0.45
	Strongly Disagree	0.12	0.13	0.75
	Strongly Agree	0.87	0.12	0.02
ASBM02G	A little Agree	0.26	0.66	0.08
	A little Disagree	0.05	0.46	0.48
	Strongly Disagree	0.11	0.12	0.77
	Strongly Agree	0.91	0.07	0.02
ASBM02H	A little Agree	0.32	0.64	0.04
	A little Disagree	0.05	0.52	0.43
	Strongly Disagree	0.13	0.11	0.77
	Strongly Agree	0.88	0.11	0.01
ASBM02I	A little Agree	0.28	0.68	0.05
	A little Disagree	0.05	0.63	0.32
	Strongly Disagree	0.06	0.18	0.76

According to Table 8, the class probability values of the three-class latent model are 0.50 for the first class (π_1^x), 0.33 for the second class (π_2^x), and 0.17 for the third class (π_3^x). To describe the characteristics of the three classes identified in the three-class latent model and interpret the behavior of the participant groups, the key parameter to be examined is the conditional probability values calculated based on respondents' answers. Examining Table 8, it is observed that the conditional probability values for the first latent class are concentrated between 0.71 and 0.91 for the "strongly agree" response. Another notable aspect of the first latent class is that the conditional probabilities for "agree a little" and "disagree a little" responses are significantly lower, ranging from 0.02 to 0.33. The answering behavior of respondents in the first latent class resembles response bias, a pattern commonly observed in affective scales (Kankaras, 2010). Another indicator of this behavior is that, despite the presence of two reverse-coded items (ASBM02B, ASBM02B) in the dataset, no change is observed in respondents' answering behavior. A detailed examination reveals that respondents in the first class exhibit an extreme response style (Davidov et al., 2014; Triandis, 1972; Van de Vijver & Leung, 2021). The conditional probability values for the second class in the latent model are concentrated on the "agree a little" and "disagree a little" responses. Based on these conditional probability values, respondents in this class tend to favor moderate options, regardless of the item content. This finding aligns with the "midpoint response style," identified by Van de Vijver and Leung (2021) as a reflection of the humility norm prevalent in East Asia. In the third latent class, respondents predominantly choose the "disagree a little" and "strongly disagree" options. This behavior does not align with any specific response style discussed in the literature. However, examining the conditional probability values for all items suggests that respondents in this class tend to exhibit negative attitudes toward learning mathematics. Based on the characteristics of the three latent classes, they have been named as follows, starting from the first class: "latent class representing extreme response tendency," "latent class representing midpoint response style," and "latent class representing participants with negative attitudes." The distribution of the three latent classes across the six countries examined in the study is presented in Figure 1.

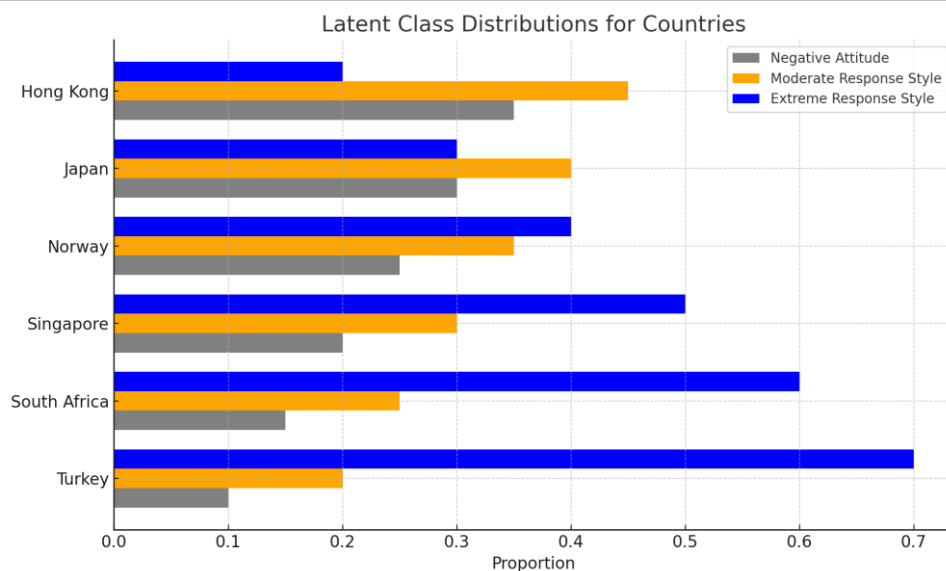


Figure 1. Latent Class Distributions for Countries Included in the Research Sample

According to Figure 1, Turkey and South Africa are predominantly positioned in the first class, representing the "extreme response style," with a probability value of approximately 0.70. Japan stands out in the latent class representing the "midpoint response style," with a probability value of around 0.50. In countries like Hong Kong, Norway, and Singapore, the distribution of respondents across the latent classes appears to be relatively more balanced.

Following the determination and explanation of the most appropriate latent class model for the data examined in the study, both statistically and theoretically, MG-LCA was applied to evaluate measurement invariance. The parameters related to the MG-LCA results for the comparison groups in which measurement invariance was analyzed across all pairwise combinations are presented in Table 9.

Table 9. Model Parameters for MG-LCA Results

Group-Model	BIC(LL)	AIC(LL)	Par.	L ²	df
1-Singapore-Hong Kong					
Heterogeneous Model	167480.84	166286.48	166	35943.88	9681
Partially Homogeneous Model	167381.83	166576.01	112	36341.41	9735
Homogeneous Model	167562.41	166950.84	85	36770.24	9762
2-Singapore-Japan					
Heterogeneous Model	177827.38	176618.59	166	35205.65	10575
Partially Homogeneous Model	177943.93	177128.37	112	35823.42	10629
Homogeneous Model	180742.49	180123.54	85	38872.61	10656
3-Singapore-Norway					
Heterogeneous Model	181128.23	179918.35	166	35453.19	10646
Partially Homogeneous Model	180940.19	180123.89	112	35766.73	10700
Homogeneous Model	182148.43	181528.91	85	37225.75	10727
4-Singapore-Turkey					
Heterogeneous Model	168273.92	167059.53	166	33199.02	10944
Partially Homogeneous Model	168169.95	167350.59	112	33598.09	10998
Homogeneous Model	170720.93	170099.11	85	36400.59	11025
5-Singapore-South Africa					
Heterogeneous Model	283043.48	281755.01	166	55857.77	17193
Partially Homogeneous Model	284540.11	283670.78	112	57881.55	17247
Homogeneous Model	290596.57	289936.81	85	64201.58	17274
6-Hong Kong-Japan					
Heterogeneous Model	123515.51	122370.11	166	28527.63	7166
Partially Homogeneous Model	123356.18	122583.38	112	28848.91	7220
Homogeneous Model	125000.69	124414.19	85	30733.71	7247
7-Hong Kong-Norway					
Heterogeneous Model	126816.86	125669.86	166	28775.16	7237
Partially Homogeneous Model	126699.02	125925.14	112	29138.45	7291
Homogeneous Model	127225.89	126638.57	85	29905.88	7318

Group-Model	BIC(LL)	AIC(LL)	Par.	L ²	df
8-Hong Kong-Turkey					
Heterogeneous Model	113964.58	112811.03	166	26520.98	7535
Partially Homogeneous Model	114011.64	113233.34	112	27051.29	7589
Homogeneous Model	115293.21	114702.54	85	28574.49	7616
9-Hong Kong-South Africa					
Heterogeneous Model	228758.68	227506.51	166	49179.74	13784
Partially Homogeneous Model	229708.49	228863.66	112	50644.89	13838
Homogeneous Model	232611.49	231970.31	85	53805.54	13865
10-Japan-Norway					
Heterogeneous Model	137167.92	136001.99	166	28036.95	8131
Partially Homogeneous Model	137156.49	136369.84	112	28512.81	8185
Homogeneous Model	139695.63	139098.62	85	31295.58	8212
11-Japan-Turkey					
Heterogeneous Model	124314.92	123143.14	166	25782.75	8429
Partially Homogeneous Model	124462.91	123672.31	112	26419.92	8483
Homogeneous Model	128174.05	127574.04	85	30375.65	8510
12-Japan-South Africa					
Heterogeneous Model	239101.09	237838.61	166	48441.49	14678
Partially Homogeneous Model	240398.19	239546.39	112	50257.28	14732
Homogeneous Model	248575.11	247928.65	85	58693.54	14759
13-Norway-Turkey					
Heterogeneous Model	127616.04	126442.89	166	26030.28	8500
Partially Homogeneous Model	127599.07	126807.55	112	26502.94	8554
Homogeneous Model	129845.96	129245.25	85	28994.64	8581
14-Norway-South Africa					
Heterogeneous Model	242401.65	241138.37	166	48689.04	14749
Partially Homogeneous Model	243787.57	242935.24	112	50593.91	14803
Homogeneous Model	249398.86	248752.00	85	56464.67	14830
15-Turkey-South Africa					
Heterogeneous Model	229546.11	228279.54	166	46434.86	15047
Partially Homogeneous Model	230404.02	229549.47	112	47812.79	15101
Homogeneous Model	232199.32	231550.78	85	49868.09	15128

The most commonly used fit indices for determining the level of measurement invariance with MG-LCA are the AIC and BIC information criteria. In MG-LCA evaluations of measurement invariance, the invariance level at which the AIC and BIC information criteria have the lowest value is considered the valid model (Magidson et al., 2020). However, simulation studies have shown that the AIC criterion provides more accurate results for small samples, while the BIC criterion is more effective for large samples (Güngör-Culha, 2012; Kankaras, 2010). Since the data analyzed in this study represent large samples, the BIC information criterion was selected as the appropriate index. Accordingly, examining Table 9, it is determined that partial homogeneity is achieved between Singapore and Hong Kong, Norway, Turkey; Hong Kong and Japan, Norway; and Norway and Turkey, Japan. For the remaining eight comparison groups, measurement invariance was found to remain at the heterogeneous model level.

DISCUSSION, CONCLUSION AND SUGGESTIONS

In this study, cross-cultural measurement invariance was examined using MG-CFA and MG-LCA techniques based on data from the Students Like Learn Mathematics Scale (SLLMS) included in the TIMSS 2019 assessment. Measurement invariance analyses were conducted across 15 comparison groups, encompassing all pairwise comparisons of the cultures of Hong Kong, South Africa, Japan, Norway, Turkey, and Singapore. The MG-CFA results for the 15 comparison groups indicated that strict measurement invariance was achieved only between Singapore-Hong Kong and Hong Kong-Norway. Measurement invariance between Turkey and South Africa was limited to the structural invariance model level, while it remained at the metric invariance level for the other 12 groups. The MG-LCA results revealed that partial homogeneity was established for measurement invariance among the cultures of Singapore-Hong Kong, Singapore-Norway, Singapore-Turkey, Hong Kong-Japan, Hong Kong-Norway, Japan-Norway, and Norway-Turkey. For the remaining 8 groups, measurement invariance remained at the heterogeneous model level. Accordingly, it was found that in MG-LCA, for the cultures where measurement invariance was achieved at the partial homogeneity level—Singapore-Norway, Singapore-Turkey, Singapore-South Africa, Hong Kong-Japan, Japan-Norway, and Norway-Turkey—measurement invariance was achieved at the metric invariance level in MG-CFA. The literature supports that the partial homogeneity model in MG-LCA corresponds to the metric invariance level in MG-CFA (Güngör-Culha, 2012; Kankaras et al., 2011; Yandı et al., 2017). The metric invariance level in MG-CFA indicates that the latent structure under evaluation shows significant similarity across different groups in terms of the number of factors, item factor loadings, and the factors to which the items belong

(Wu et al., 2007). More broadly, the metric invariance model suggests that the regression curves representing the relationship between indicators and the latent variable are equivalent (Kline, 2016). Examining the practical implications of the metric invariance model, Steenkamp and Baumgartner (1998) found that respondents from different groups with metric invariance tend to respond similarly to items or create similar response patterns. In the partial homogeneity model in MG-LCA, it is assumed that the relationship between latent classes and items is consistent across comparison groups. However, the partial homogeneity model does not constrain class distributions and item distributions to remain constant. This implies that latent class structures are similar across comparison groups, but respondent patterns may vary (Kankaras, 2010). In measurement invariance research, the findings obtained from statistical analyses must be examined and explained from both theoretical and practical perspectives. From this viewpoint, the MG-CFA results showing strict measurement invariance between the cultures of Singapore and Hong Kong is an expected outcome. Factors such as the geographical proximity of the two countries, the preference for English as the medium of instruction in both, similarities in educational procedures, and cultural similarities are thought to contribute to achieving strict invariance (Hambleton, 2004; Maden-Kalkan & Yılmaz-Şaşmaz, 2021; Özçelik-Tezel, 2007). An unexpected finding is the establishment of strict measurement invariance between Hong Kong and Norway. Although Hong Kong and Norway are influenced by geographically distinct cultures, such as Northern Europe and Asia, this strict invariance may stem from their shared emphasis on early childhood education, similarities in educational frameworks, and comparable levels of social development (Foy & LaRoche, 2020; Küçüköğlu & Ercan, 2019). The MG-CFA results indicated that Japan achieved metric invariance with all other countries. This suggests that the latent structure examined in the study is similar in Japan and other cultures in terms of the number of factors, the distribution of items across factors, and factor loadings (Kline, 2016). The comparison group with the lowest level of measurement invariance was between Turkey and South Africa, where only structural invariance was achieved. This implies that the latent structure examined in the study shares the same number of factors for Turkey and South Africa, but factor loadings may differ (Kline, 2016).

In addition to MG-LCA, detailed insights into the classes formed by the cultures in the study sample were obtained through latent class analysis (LCA). According to the findings, participants from South Africa, Turkey, and Japan exhibited response patterns similar to certain response styles discussed in the literature. Existing literature emphasizes that response styles substantially impact measurement invariance, particularly in cross-cultural comparisons (Kankaras, 2010; Van de Vijver & Leung, 2021). The MG-LCA results showed that in the comparison groups examined in the study, when both cultures exhibited response styles, there was a 100% likelihood of the heterogeneous model being accepted. Conversely, when neither group exhibited response styles, the partial homogeneity model was accepted with a 100% likelihood. When only one group exhibited a response style, the partial homogeneity model was deemed compatible with a 45% likelihood. From the perspective of MG-CFA, when both countries in the comparison group exhibited response styles, the metric invariance model was identified as appropriate with a 67% likelihood, and the structural invariance model with a 33% likelihood. If only one of the two countries exhibited a response style, the metric invariance model was found to be appropriate with a 100% likelihood. If neither country exhibited response styles, there was a 67% likelihood that the strict invariance model was appropriate.

Overall, the results indicate that when MG-CFA and MG-LCA are used together in cross-cultural studies, they approach measurement invariance from different perspectives and, to some extent, complement each other. It is believed that using these two methods together, rather than relying on a single statistical method, provides more comprehensive insights into questions such as "why measurement invariance could not be achieved." Hambleton (2004) and Borsboom (2006) emphasized that in measurement invariance studies, researchers should not merely determine the statistical level of invariance, as this is an incomplete approach. Instead, the primary objective should be to identify "the reasons why measurement invariance could not be achieved." From this perspective, it was found that response styles identified through latent class analysis approaches also align with the results of MG-CFA. It is suggested that repeating the application of MG-LCA and MG-CFA with different latent traits, samples, and conditions by various researchers would contribute significantly to the field of measurement invariance research.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, author-ship, and/or publication of this article.

Statements of publication ethics

We hereby declare that the study has not unethical issues and that research and publication ethics have been observed carefully.

Researchers' contribution rate

First author conceived of the presented idea and developed the theory. First author performed the computations and conducted the analyses. Second author verified the analytical methods and supervised the research process. Second author also provided critical feedback and guided First author in structuring the study and interpreting the results. Both authors discussed the findings and contributed to the final manuscript.

Ethics Committee Approval Information

The ethics approval for this study was granted by the Ankara University Ethics Committee on August 2, 2021, with approval number 13/276 and reference number 85434274-050.04.04 / 226918.

REFERENCES

- Altıntaş, Ö. (2016). *Ankara Üniversitesi yabancı uyruklu öğrenci seçme testinin ölçme değişmezliğinin örtük sınıf ve Rash Modeline göre incelenmesi* (Yayınlanmamış Doktora Tezi), Ankara Üniversitesi, Ankara.
- Arıcıgil-Çılan, Ç. (2015). *Uygulamalı gizli sınıf analizi*. İstanbul: Çağlayan Kitabevi.
- Bergman, L.R., & Magnusson, D. (1997). A person-oriented approach in research on developmental psychopathology. *Development and psychopathology*, 9(2), 291-319.
- Bergman, L.R., & Wangby, M. (2014). The person-oriented approach: A short theoretical and practical guide. *Estonian Journal of Education*, 2(1), 29-49. doi: <http://dx.doi.org/10.12697/eha.2014.2.1.02b>
- Borsboom, D. (2006). When does measurement invariance matter? *Medical care*, 44(11), 176-181.
- Brown, T. A., & Moore, M. T. (2012). Confirmatory factor analysis. In R. H. Hoyle (Ed.), *Handbook of structural equation modeling* (pp. 361-379). New York, NY: Guilford Press.
- Cheung, G.W., & Rensvold, R.B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling: A Multidisciplinary Journal*, 9(2), 233-255. doi: http://dx.doi.org/10.1207/S15328007SEM0902_5
- Cieciuch, J., Davidov, E., Schmidt, P., & Algesheimer, R. (2019). How to obtain comparable measures for cross-national comparisons. *Kölner Zeitschrift für Soziologie und Sozialpsychologie*, 71(S1), 157-186. <https://doi.org/10.5167/uzh-172539>
- Clogg, C.C., & Goodman, L.A. (1985). Simultaneous latent structure analysis in several groups. *Sociological Methodology*, 15(1), 81-110.
- Collins, L. M., & Lanza, S. T. (2009). *Latent class and latent transition analysis: With applications in the social, behavioral, and health sciences*. Hoboken, NJ: Wiley. <https://doi.org/10.1002/9780470567333>
- Coşkun, F. (2023). PISA 2018 verisi üzerinden öğrencilerin kültürlerarası iletişim algısının örtük sınıf modelleri ile incelenmesi: türkiye örnekleme. *Gazi Eğitim Bilimleri Dergisi*, 9(1), 1-20.
- Coşkun, F., & Gülleroğlu, H. D. (2023). Öğrencileri eğitimde internet kullanımına yönlendirmeye ilişkin öğretmen profillerinin PISA 2018 uygulaması kapsamında karşılaştırması. *Journal of History School*, 16(LXV), 1530-1556. <https://doi.org/10.29228/joh.69740>
- Çüm, S., Demir, E. K., Akın Arıkan, Ç., & Şahin, M. D. (2020). Okuma Zevki Ve Okuma Çeşitliliğinin Örtük Sınıf Analizi: Türkiye Ve Çin İncelemesi. *Gazi Üniversitesi Gazi Eğitim Fakültesi Dergisi*, 40(3), 943-977.
- Davidov, E., Meuleman, B., Cieciuch, J., Schmidt, P., and Billiet, J. (2014). Measurement equivalence in cross-national research. *Annual Review of Sociology*, 40(1), 55-75. doi: <http://dx.doi.org/10.1146/annurev-soc-071913-043137>
- Demir, E. (2019). *R diliyle istatistik uygulamaları*. Ankara: Pegem Akademi.
- DeMars, C.E. (2016). Partially compensatory multidimensional item response theory models: Two alternate model forms. *Educational and Psychological Measurement*, 76(2), 1-27. doi: <http://dx.doi.org/10.1177/0013164415589595>
- Eid, M., Langeheine, R., & Diener, E. (2003). Comparing typological structures across cultures by multigroup latent class analysis: A primer. *Journal of Cross-Cultural Psychology*, 34(2), 195-210. <https://doi.org/10.1177/0022022102250427>
- Ercikan, K., Por, H.-H., & Guo, H. (2023). Cross-cultural validity and comparability in assessments of complex constructs. *Educational Testing Service*. <https://oecd-ilibrary.org>
- Foy, P. & LaRoche, S. (2020). Estimating standard errors in the TIMSS 2019 results. In M.O. Martin, M. von Davier, and I.V.S. Mullis, (Eds.) *TIMSS 2019 Technical report*. <https://timssandpirls.bc.edu/timss2019/methods>
- Fraenkel, J.R., & Wallen, N. E. (2009). *How to design and evaluate research in education* (7th ed.). Boston: McGraw Hill Higher Education.
- Gliner, J.A., Morgan, G.A., & Leech, N.L. (2017). *Research methods in applied settings: An integrated approach to design and analysis* (Third Edition). New York, London: Routledge Taylor & Francis Group
- Green B.F. (1952). Latent structure analysis and its relation to factor analysis. *Journal of the American Statistical Association*, 47(257), 71-76.
- Güngör, D., Korkmaz, M., & Somer, O. (2013). Çoklu-grup örtük sınıf analizi ve ölçme eşdeğerliği. *Türk Psikoloji Dergisi*, 28(72), 48-57.
- Güngör-Culha, D. (2012). *Örtük sınıf analizlerinde ölçme eşdeğerliğinin incelenmesi* (Yayınlanmamış doktora tezi). Ege Üniversitesi, İzmir.
- Hambleton, R. K., & Zenisky, A. L. (2011). Translating and adapting tests for cross-cultural assessments. In R. K. Hambleton, P. F. Merenda, & C. D. Spielberger (Eds.), *Adapting educational and psychological tests for cross-cultural assessment* (pp. 3-38). Mahwah, NJ: Erlbaum.
- Hambleton, R.K. (2004). Issues, designs, and technical guidelines for adapting tests into multiple languages and cultures. R.K. Hambleton, P.F. Merenda and C.D. Spielberger (Eds). *Adapting educational and psychological tests for cross-cultural assessment* (pp. 3-40). London: Psychology Press.
- Hernández-Torrano, D., & Courtney, M. G. R. (2021). Modern international large-scale assessment in education: An integrative review and mapping of the literature. *Large-scale Assessments in Education*, 9(17). <https://doi.org/10.1186/s40536-021-00109-1>
- Janousch, C., Sinha, M., & Springer, M. (2022). Resilience profiles across context: A latent profile analysis in a German, Greek, and Swiss sample of adolescents. *PLOS ONE*, 17(1), e0262039. <https://doi.org/10.1371/journal.pone.0262039>
- Johansson, S. (2016). International large-scale assessments: What uses, what consequences? *Educational Research*, 58(2), 139-148. <https://doi.org/10.1080/00131881.2016.1165559>
- Johnson, T. (1998). Approaches to equivalence in cross-cultural and cross-national survey research. In J. A. Harkness (Ed.), *Zuma-Nachrichten Spezial Volume 3: Cross-Cultural Survey Equivalence* (pp. 1-40). Mannheim: Zuma.
- Kankaraş M., Moors G., and Vermunt J.K. (2011). Testing for measurement invariance with latent class analysis. In E. Davidov, P. Schmidt, J. Billiet (Eds.), *Cross-cultural analysis: Methods and applications* (pp. 359-384). New York: Routledge.
- Kankaras, M. (2010). *Essays on measurement equivalence in cross-cultural survey research: A latent class approach* (Unpublished PhD thesis). Tilburg University, Netherlands.
- Kankaras, M., & Vermunt, J.K. (2014). Simultaneous latent class analysis across groups. A.C. Michalos (Ed.) *Encyclopedia of quality of life and well-being research* (pp. 5969-5974). Heidelberg: Springer.
- Kankaraş, M., Moors, G., & Vermunt, J. K. (2018). Testing for measurement invariance with latent class analysis. In E. Davidov, P. Schmidt, & J. Billiet (Eds.), *Cross-cultural analysis* (2nd ed., pp. 361-384). Routledge. <https://doi.org/10.4324/9781315537078>
- Kankaras, M., Vermunt, J. K., & Moors, G. (2011). Measurement equivalence of ordinal items: A comparison of factor analytic, item response theory, and latent class approaches. *Sociological Methods & Research*, 40(2), 279-310. <https://doi.org/10.1177/0049124111405301>

- Kim, E. S., Cao, C., Wang, Y., & Nguyen, D. T. (2017). Measurement invariance testing with many groups: A comparison of five approaches. *Structural Equation Modeling: A Multidisciplinary Journal*, 24(4), 524-544. <https://doi.org/10.1080/10705511.2017.1304822>
- Kline, R.B. (2016). *Principles and practice of structural equation modeling* (Fourth Edition). New York, London: The Guilford Press.
- Kirsch, I., & Braun, H. (2020). Changing times, changing needs: Enhancing the utility of international large-scale assessments. *Large-scale Assessments in Education*, 8(10). <https://doi.org/10.1186/s40536-020-00088-9>
- Küçüköğlu, M., & Ercan, H. (2019). Norveç'te refah devletinin ortaya çıkışı ve gelişimi. *Uluslararası Toplum Araştırmaları Dergisi*, 11(18), 2275-2308. <https://doi.org/10.26466/opus.501680>
- Lin, H. T. (2006). A comparison of model selection indices for nested latent class models. *Monte Carlo Methods and Applications*, 12(3), 239-259. <https://doi.org/10.1515/156939606778705164>
- Lubke, G., & Neale, M. (2008). Distinguishing between latent classes and continuous factors with categorical outcomes: Class invariance of parameters of factor mixture models. *Multivariate Behavioral Research*, 43(4), 592-620. <https://doi.org/10.1080/00273170802490673>
- Maden-Kalkan, Ç., & Yılmaz-Şaşmaz, A. (2021). Covid-19 Salgınının Kentsel Yaşam Kalitesi Açısından Potansiyel Etkileri Çin Örneği. *Kent Akademisi*, 14(4), 1283-1298.
- Magidson, J., & Vermunt, J.K. (2004). Latent class models. In D. Kaplan (Ed.), *Handbook of quantitative methodology for the social sciences* (pp. 345-368). Newbury Park, CA: Sage Publications.
- Magidson, J., Vermunt, J.K., & Madura, J. P. (2020). *Latent class analysis*. London: SAGE Publications Limited.
- Meade, A. W., Johnson, E. C., & Braddy, P. W. (2008). Power and sensitivity of alternative fit indices in tests of measurement invariance. *Journal of applied psychology*, 93(3), 568-592. <https://doi.org/10.1037/0021-9010.93.3.568>
- Meitinger, K., Davidov, E., Schmidt, P., & Braun, M. (2020). Measurement invariance: Testing for it and explaining why it is absent. *Survey Research Methods*, 14(4), 345-349. <https://doi.org/10.5167/uzh-192239>
- Meredith, W. (1993). Measurement invariance, factor analysis and factorial invariance. *Psychometrika*, 58(4), 525-543.
- Mullis, I.V.S., Martin, M.O., Foy, P., Kelly, D. L., & Fishbein, B. (2020). *TIMSS 2019 International Results in Mathematics and Science*. Retrieved from Boston College, TIMSS & PIRLS International Study Center website: <https://timssandpirls.bc.edu/timss2019/international-results/>
- Oberski, D. (2016). Mixture models: Latent profile and latent class analysis. In J. Robertson and M. Kaptein (Eds.), *Modern Statistical Methods for HCI*. Cham, Switzerland: Springer International Publishing.
- Özcelik-Tezel, C. (2007). Ulusal kimliğin oluşumunda müze ve toplum ilişkisi: Singapur. *Elektronik Sosyal Bilimler Dergisi*, 6(20), 133-155.
- Steenkamp, J.B.E., & Baumgartner, H. (1998). Assessing measurement invariance in cross-national consumer research. *Journal of consumer research*, 25(1), 78-90.
- Tabachnick, B.G., & Fidell, L.S. (2019). *Using multivariate statistics* (Seventh Editon). Boston, MA: Pearson.
- Thissen, D. (2001). *IRTLRDIF v. 2.0 b: Software for the computation of the statistics involved in item response theory likelihood-ratio tests for differential item functioning*. Chapel Hill, NC: LL Thurstone Psychometric Laboratory.
- Triandis, H. C. (1972). *The analysis of subjective culture*. New York, NY: Wiley.
- Van De Schoot, R., Schmidt, P., De Beuckelaer, A., Lek, K., and Zondervan-Zwijnenburg, M. (2015). Measurement invariance. *Frontiers in psychology*, 6(1064). doi: <https://10.3389/fpsyg.2015.01064>
- Van de Vijver, F.J., & Leung, K. (2021). *Methods and data analysis for cross-cultural research* (Second Education). United Kingdom: Cambridge University Press.
- Van de Vijver, F.J.R. (1998). Towards a theory of bias and equivalence. J.A. Harkness (Ed.), Zuma-Nachrichten Spezial Volume 3. *Cross-Cultural Survey Equivalence* (pp. 41-65). Mannheim: Zuma.
- Van de Vijver, F.J.R., & Leung, K. (2011). Equivalence and bias: A review of concepts, models, and data analytic procedures. D. Matsumoto and F.J.R. Van de Vijver (Eds.), *Cross-cultural research methods in psychology* (pp. 17-45). New York, NY: Cambridge University Press.
- Vandenberg, R.J., and 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.
- Wendt, H., Kasper, D., & Trendtel, M. (2017). Assuming measurement invariance of background indicators in international comparative educational achievement studies: A challenge for the interpretation of achievement differences. *Large-scale Assessments in Education*, 5(10). <https://doi.org/10.1186/s40536-017-0043-9>
- Wu, A.D., Li, Z., and 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. *Practical Assessment, Research, and Evaluation*, 12(3), 1-26. <https://doi.org/10.7275/mhqa-cd89>
- Yandı, A., Köse, İ.A., ve Uysal, Ö. (2017). Farklı yöntemlerle ölçme değişmezliğinin incelenmesi: PISA 2012 örneği. *Mersin Üniversitesi Eğitim Fakültesi Dergisi*, 13(1), 243-253.
- Yin, L., and Fishbein, B. (2020). Creating and interpreting the TIMSS 2019 context questionnaire scales. In M.O. Martin, M. von Davier, & I.V.S. Mullis (Eds.), *Methods and Procedures: TIMSS 2019 Technical Report* (pp. 16.1-16.331). Boston College, TIMSS & PIRLS International Study Center.
- Zhao, M., & Jin, R. (2023). Advancing a cross-cultural understanding of teacher perceptions of school climate: A latent class analysis using 2018 TALIS data. *Frontiers in Psychology*, 14, 1129306. <https://doi.org/10.3389/fpsyg.2023.1129306>



Development of the Family Education Activities Evaluation Scale

Aile Eğitim Etkinliklerinin Değerlendirilmesi Ölçeğini Geliştirme Çalışması¹

Sinem SERÇE², Naim ÜNVER³

Keywords

1. Scale development
2. Family education evaluation
3. Attitude
4. Communication
5. Expectation

Anahtar Kelimeler

1. Ölçek geliştirme
2. Aile eğitimi değerlendirme
3. Tutum
4. İletişim
5. Beklenti

Received/Başvuru Tarihi

27.03.2025

Accepted / Kabul Tarihi

10.10.2025

Abstract

Purpose: This study aimed to revise the "Evaluation Scale of OBADER Family Education Sessions Delivered via Distance Education," developed within a thesis, into a more general scale titled the "Family Education Activities Evaluation Scale."

Design/Methodology/Approach: The study employed the survey model, a quantitative research method, and was conducted in two phases. In the first phase, the original version of the scale was developed based on data collected from 310 participants. In the second phase, carried out two years later, the revised version of the scale was administered to a new sample of 241 participants.

Findings: In both phases, the scale consisted of 16 items grouped under 3 factors. While the item contents were slightly modified during the revision process, the overall factor structure was preserved. The results confirmed the scale's validity and reliability across both samples.

Highlights: The final version of the scale, titled "Family Education Activities Evaluation Scale," provides a valid and reliable tool for evaluating various family education activities beyond the initial OBADER context, making it adaptable for broader applications in distance education and family training programs.

Öz

Çalışmanın amacı: Bu çalışmanın amacı, piyano eğitiminde üstbilişsel stratejilerin kullanımının üstbilişsel farkındalık ve piyano eğitimi üzerindeki etkilerini incelemek ve bu doğrultuda geliştirilen "OBADER Aile Eğitimi Etkinliklerinin Değerlendirilmesi Ölçeği"nin daha genel bir ölçek haline dönüştürülmesini amaçlamaktadır.

Materyal ve Yöntem: Araştırmada nicel araştırma yöntemlerinden biri olan tarama modeli kullanılmıştır ve çalışma iki aşamalı olarak yürütülmüştür. İlk aşamada ölçek geliştirme çalışması 310 katılımcıdan elde edilen verilerle gerçekleştirilmiştir. İkinci aşama ise iki yıl sonra farklı bir örneklem grubu olan 241 katılımcı ile yürütülmüştür.

Bulgular: Her iki aşamada da ölçek 3 faktör altında toplanan 16 maddeden oluşmuştur. Ölçekteki bazı madde ifadeleri ikinci aşamada revize edilmiş olsa da faktör yapısı korunmuştur. Elde edilen sonuçlar ölçeğin geçerlik ve güvenirlik açısından tutarlı olduğunu göstermiştir.

Önemli Vurgular: Son haliyle "Aile Eğitimi Etkinliklerini Değerlendirme Ölçeği" olarak adlandırılan ölçek, sadece OBADER bağlamı için değil, farklı uzaktan eğitim ve aile eğitimi etkinliklerinde de kullanılabilecek geçerli ve güvenilir bir ölçek olarak öne çıkmaktadır.

¹ This study is derived from the master's thesis titled "Investigation of the Effect of OBADER Sessions Given By Distance Education on Parents With 36-72 Months Old Children."

² Kastamonu, Türkiye; <https://orcid.org/0000-0003-2937-1764>

³ Kastamonu Üniversitesi, Eğitim Fakültesi, Okul Öncesi Eğitimi ABD, Kastamonu, Türkiye, <https://orcid.org/0000-0003-1801-5609>

INTRODUCTION

As one of the fundamental components of early childhood education, the family has long been a significant focus of research, and family education continues to be a crucial field of study today (2025), just as it was in the past. The efforts of high-quality early childhood education programs to establish strong partnerships with families stem from the fact that parents play a vital role in supporting their children's social competence, self-regulation skills, and school readiness during the preschool years (Slotkin et al., 2024). The active involvement of parents in their children's educational process is considered highly important worldwide and is increasingly viewed as an integral part of school systems. In Turkey as well, with recent reforms in the field of education, the active participation of parents in the educational process is being encouraged. Through various institutions, structured and systematic programs are provided for parents on topics such as child development and education, as well as the elimination of negative behaviors (Özyürek et al., 2015).

A review of the literature reveals a variety of definitions related to family education. Family education encompasses activities aimed at providing families with systematic guidance to help them support their children's development and actively participate in their educational journey, ultimately fostering the skills necessary to raise healthy children/individuals (Tezel Şahin & Özyürek, 2008; Mahoney et al., 1999). "Parent education," a term often used interchangeably, refers to organized efforts with clearly defined content, target audiences, and goals designed to improve or change parental role performance, ultimately facilitating parenting behaviors that positively influence child development outcomes. Parent education is grounded in the principles that parenting is a complex, challenging, yet rewarding responsibility, and that education can help parents fulfill their roles more effectively (Smith, Perou, & Lesene, 2002).

Family education programs should be delivered by professionals, taking into consideration the sociocultural characteristics of the participants. Communication with families should be respectful, clear, and aligned with human rights principles. Patience, honesty, and fairness must be demonstrated throughout the activities, and educators are expected to provide guidance. When organizing family education activities, several key principles should be observed: scientific validity, utility, cooperation, planning, confidentiality, voluntariness, transparency, accessibility, continuity, interactivity, participation, and appropriateness (Yılmaz Bolat, 2017; MoNE, 2013b; Ünal, 2018).

An examination of the history of preschool education programs in Turkey shows that the 1953 curriculum introduced the concept of school-family cooperation under the heading "Relations with Parents and the Community." The 1989 Preschool Education Program emphasized the need to plan parental involvement in educational activities. The Ministry of National Education (MoNE) 1994 curriculum highlighted the importance of family involvement for the permanence of education and recommended conducting interviews with parents. The 2002 Preschool Education Program elaborated in detail on the need to ensure family participation in children's education. In the 2006 curriculum, family education activities were defined as structured and planned efforts aimed at enhancing parents' knowledge and skills regarding child health, development, behavior management, communication, nutrition, and mental health. The program also outlined several tools and methods for implementing family education—such as meetings, conferences, individual consultations, printed materials like articles, brochures, handbooks, and educational boards—and introduced the "Family Education Needs Assessment Form." Teachers were advised to analyze these forms completed by families to plan education activities accordingly, a recommendation that continued in subsequent programs.

The 2013 curriculum updates included two separate age groups: "The Education Program for Children Aged 0–36 Months" and "The Preschool Education Program for Children Aged 36–72 Months." Alongside these programs, two comprehensive family support education guides were developed: EBADER (Family Support Education Guide Integrated with the 0–36 Month Program) and OBADER (Family Support Education Guide Integrated with the 36–72 Month Preschool Education Program), underscoring the importance placed on family education and participation. In the 2024 update of the Preschool Education Program, a new "Family Education Guide" was prepared. The 2024 Century of Türkiye Maarif Model Preschool Education Program describes family education activities as practices designed to meet families' needs by enhancing and supporting their child-rearing skills. In this program, planning family education activities for families with preschool-aged children is listed among the "school and community participation" activities. To ensure quality planning and analysis of these activities, the program includes the "Family Education Needs Assessment Form." While OBADER recommends face-to-face delivery of family education activities, the Century of Türkiye Maarif Model allows for both face-to-face and online modalities (MoNE, 2002, 2006, 2013a, 2013b, 2013c, 2024a, 2024b).

OBADER is a guide launched in 2013, developed for families with children aged 3–6, containing components related to family education and participation. Designed to facilitate cooperation between schools and families, OBADER aims to promote active involvement of parents in both family education and classroom activities. It also encourages the participation of parents of preschool-aged children in the educational process. Furthermore, OBADER serves as a resource for educators, aiming to enhance their knowledge and awareness of the family education and participation process. It guides them in how to plan, implement, manage, and evaluate this process (MoNE, 2013b; Ünlü Çetin, 2016).

Distance education is a learning environment where learners and educators interact through communication and educational technologies without constraints of time or location. It provides learners with the opportunity to study independently, offering individual flexibility while eliminating the requirement for physical presence. One of the key advantages of distance education is its capacity to minimize spatial and temporal limitations, thereby ensuring equal access to educational opportunities. It also offers various alternative learning options such as technology-based learning, open education, and virtual applications (Uşun, 2006;

Koçak & Sallabaş, 2022). Based on these characteristics, the current study focuses on adapting OBADER family education activities, originally intended for face-to-face delivery, to an online distance education format.

Purpose of the Study

The purpose of this study is to revise the Evaluation Scale of OBADER Family Education Sessions Delivered via Distance Education (UEOAEODÖ) into the Evaluation Scale of Family Education Activities in Preschool (OAEDÖ). This revision was deemed necessary due to the discontinuation of the OBADER framework.

METHOD AND FINDINGS

In this study, data were collected using the survey model, one of the quantitative research designs. The survey model is a research approach aimed at gathering data to determine the specific characteristics of a particular group (Büyükoztürk et al., 2020).

1. Development of the Data Collection Tool Data Collection

Ensuring that the scale items are clearly and accurately understood is crucial for obtaining valid responses during administration. Moreover, the instructions provided within the scale must clearly and understandably convey the purpose, content, and completion process of the scale. The clarity and comprehensibility of the instructions play a critical role in the application and interpretation of the scale (Koçak & Sallabaş, 2022).

In the development process of the UEOAEODÖ, a 60-item pool was created based on a literature review and expert opinions to ensure content validity. A preliminary study revealed that 10 of these items were not clearly understood by participants and were subsequently removed, resulting in a final pool of 50 items.

2. Data Collection

The final version of the 50-item pool was administered to 350 parents. Upon evaluating the control questions, 40 responses were deemed unreliable and were excluded from the analysis. The data from the remaining 310 participants were used for factor analysis. The scale was based on a five-point Likert scale:

- 1 = Strongly Disagree,
- 2 = Mostly Disagree,
- 3 = Somewhat Agree,
- 4 = Mostly Agree,
- 5 = Strongly Agree.

Initially, Exploratory Factor Analysis (EFA) was conducted to identify factor structures, followed by Confirmatory Factor Analysis (CFA) to validate the structure, leading to the development of the UEOAEODÖ.

Due to the discontinuation of the OBADER framework, which was implemented under the 2013 Preschool Education Program by the Turkish Ministry of National Education and ceased to be effective in 2024, the UEOAEODÖ was revised. Modifications were made to 16 items and some Likert-scale expressions. Subsequently, data were collected from a different sample of 270 parents. Among these, 15 were excluded due to careless responses. A Z-score analysis was applied to detect outliers among the remaining 255 responses, resulting in the exclusion of 4 additional responses. Confirmatory Factor Analysis was conducted on the final sample of 241 participants using AMOS software, leading to the development of the OAEDÖ.

Regarding the sample size, Nunnally & Bernstein (1994) and Can (2013) recommend that the sample size should be at least 10 times the number of items. If the total sample is below 300, it should be 5 to 10 times the number of items. Since the revised scale includes 16 items, a sample size of at least 160 is considered adequate.

All data were collected via Google Forms and shared with parents using a convenience sampling method. The form link was distributed through social media platforms (e.g., Instagram, WhatsApp), targeting parents of children aged 36–72 months. In the OAEDÖ data collection phase, a five-point Likert scale was used with the following values:

- 1 = Strongly Disagree,
- 2 = Somewhat Disagree,
- 3 = Neither Agree nor Disagree,
- 4 = Mostly Agree,
- 5 = Strongly Agree.

3. Data Analysis

The data analysis in this section pertains to the evaluation of the UEOAEODÖ developed within the master's thesis study.

Table 1. Sample group

Variables	First Sample Group	N
Parent Gender	Female	264
	Male	46
Parent Age	20–25	7
	26–30	79
	31–35	152
	36–40	50
	41–45	19
	46–50	2
	51 and above	1
Parent Educational Level	Primary School	7
	Middle School	6
	High School	26
	Associate Degree	36
	Bachelor's Degree	198
	Postgraduate	37
Parent Occupation	Not Employed	100
	Public Servant	115
	Worker	61
	Tradesperson	34

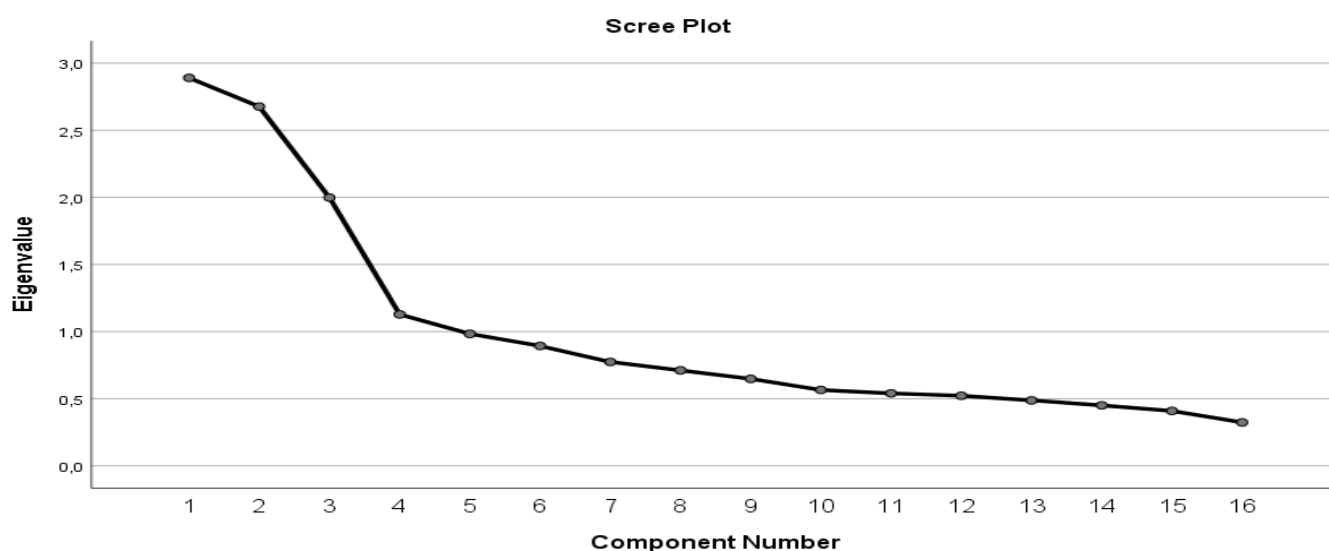
Descriptive Statistics of the First Sample Group

Among the 310 parents in the first sample group, 85.2% were female and 14.8% were male. The highest proportion of parents (49.0%) were between the ages of 31–35, while the lowest proportion (0.3%) were aged 51 and above. In terms of educational background, the largest group held a bachelor's degree (63.9%), while the smallest group had completed middle school (1.9%). Regarding occupation, the majority were public servants (37.1%), whereas the smallest group were tradespeople (10.9%).

Exploratory Factor Analysis (EFA)

In the initial EFA, the anti-image correlation matrix diagonal values were examined. Based on this analysis, four items with values close to 0.50 were removed from the scale, as they were considered to potentially have a significant negative impact on factor loading. After removing these items, the factor analysis was repeated.

As a result of the subsequent analyses, a scale consisting of 16 items and 3 factors was developed. The final EFA results showed that the scale explained 47.278% of the total variance. The corresponding scree plot for this analysis is presented in Figure 1.

**Figure 1. Scree Plot**

Upon examining the scree plot, it is observed that the curve begins to level off after the third factor. This indicates that a three-factor solution is appropriate and that the development of a three-factor scale is supported by the analysis results.

Table 2. Total Variance Explained by the Factors

Item	Initial Eigenvalues			Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings		
	Total	Variance %	Cumulative %	Total	Variance %	Cumulative %	Total	Variance %	Cumulative %
1	2.890	18.063	18.063	2.890	18.063	18.063	2.868	17.922	17.922
2	2.676	16.725	34.788	2.676	16.725	34.788	2.538	15.862	33.784
3	1.998	12.490	47.278	1.998	12.490	47.278	2.159	13.494	47.278
4	1.128	7.051	54.330						
5	0.983	6.144	60.474						
6	0.893	5.582	66.056						

The percentage of variance explained by the first factor was found to be 18.063%, by the second factor 16.725%, and by the third factor 12.490%. The cumulative variance explained by all three factors was determined to be 47.278%.

Table 3. Distribution of Items by Factors

Item	Factor Component		
	1	2	3
AB1. Preschool education should meet the developmental needs of the child.	,790		
AB2. In preschool education, parents should be regularly informed by the teacher about their child's development.	,752		
AB3. Preschool education supports my child's motivation to learn.	,694		
AB4. Preschool education should be appropriate to the child's developmental characteristics.	,659		
AB5. Preschool education prepares my child for primary school.	,634		
AB6. Preschool education supports my child in speaking Turkish correctly and fluently.	,576		
AT1. It is important for me that my child always does their best and excels.		,669	
AT2. I generally fulfill everything my child wants.		,665	
AT3. When choosing toys, I buy whatever my child wants.		,644	
AT4. I usually intervene and act controlling in everything my child does.		,641	
AT5. I buy a new toy for my child whenever they ask.		,619	
AT6. I usually protect and guard my child in everything they do.		,617	
AÇİ1. When my child expresses emotions, I make them feel understood.			,795
AÇİ2. I give my child the opportunity to better express their emotions and thoughts.			,792
AÇİ3. I choose words appropriate to my child's age and development when speaking.			,655
AÇİ4. I make eye contact when speaking with my child.			,584

As shown in the table, the factor loadings of the six items in the first factor ranged from .576 to .790. In the second factor, the factor loadings of the six items ranged from .617 to .669. For the third factor, the factor loadings of the four items ranged from .584 to .795.

Reliability Coefficient

Under this heading, the reliability of the scale was evaluated using the Cronbach's Alpha coefficient.

Table 4. Reliability Analysis of the Scale

Factors	Mean ($\bar{x}$)	SD	Variance	Number of Items	Cronbach's Alpha
Entire Scale	63.67	5.77	33.37	16	,622
Parental Expectations	28.97	2.10	4.42	6	,713
Parental Attitude	15.13	5.38	28.95	6	,719
Parent-Child Communication	19.56	1.01	1.017	4	,659

As a result of the factor analysis, the Cronbach's Alpha coefficient for the 16 remaining items was found to be .622. The closer this value is to 1, the more reliable the scale is considered to be (Can, 2013). The Cronbach's Alpha coefficient for the first factor was .713 and included six items. The second factor had a Cronbach's Alpha value of .719, also consisting of six items. The third factor had a Cronbach's Alpha of .659 and included four items. An alpha value above .75 generally indicates a highly reliable scale, while values between .50 and .75 typically suggest moderate reliability. Values below this range are usually considered to indicate low reliability (Hinton et al., 2004).

Based on the factor analysis, a three-factor scale consisting of 16 items was developed. The first factor, which includes six items, was named “parental expectations” (AB); the second factor, also with six items, was labeled “parental attitude” (AT); and the third factor, composed of four items, was named “parent-child communication” (AÇİ).

Item Analysis of the Total Scale

Under this heading, the correlation of each item with the overall scale was calculated, as well as the Cronbach’s Alpha coefficient that would result if each item were deleted from the scale.

Table 5. Item-Total Correlation of the Scale

Item	Mean if Item Deleted	Variance if Item Deleted	Item-Total Correlation	Cronbach’s Alpha if Item Deleted
AB1	58.7742	32.887	,055	,624
AB2	58.7258	32.737	,135	,620
AB3	59.0258	30.821	,220	,609
AB4	58.9935	31.469	,158	,617
AB5	58.7419	32.425	,232	,615
AB6	58.8065	32.331	,174	,616
AT1	61.0645	25.601	,398	,574
AT2	60.3258	25.864	,370	,581
AT3	61.3839	25.791	,387	,577
AT4	61.0097	23.770	,465	,555
AT5	58.7290	33.363	-,016	,626
AT6	58.7774	33.423	-,042	,629
AÇİ1	58.8677	33.565	-,076	,635
AÇİ2	58.7581	33.032	,064	,623
AÇİ3	61.0323	25.876	,412	,571
AÇİ4	62.0968	28.178	,346	,587

As shown in the overall reliability analysis table, the Cronbach’s Alpha coefficient for the entire scale was calculated as .622. It was observed that removing item AÇİ1 from the third factor would increase the reliability coefficient to .635. However, since the increase of .013 is not considered significant, and given that this item has a factor loading of .655, it was decided—based on expert opinion—not to remove the item from the scale.

Inter-Item Correlation

Table 6. Correlation Between Scale Items

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1-AB1	1,000															
2-AB2	,523	1,000														
3-AB3	,303	,418	1,000													
4-AB4	,203	,253	,366	1,000												
5-AB5	,405	,568	,264	,300	1,000											
6-AB6	,283	,397	,309	,377	,498	1,000										
7-AT1	-,107	-,069	,030	,052	-,001	-,021	1,000									
8-AT2	-,091	-,069	,051	-,019	-,012	,038	,288	1,000								
9-AT3	-,133	-,056	,086	,002	-,014	-,038	,282	,340	1,000							
10-AT4	-,066	,028	,139	,092	,071	,044	,305	,377	,471	1,000						
11-AÇİ1	-,022	,047	,018	-,027	,038	,090	-,046	-,040	-,071	-,106	1,000					
12-AÇİ2	-,004	-,049	-,104	-,095	,088	-,005	-,059	-,082	-,096	-,119	,403	1,000				
13-AÇİ3	,029	-,059	,048	,067	-,001	,061	-,110	-,134	-,157	-,111	,161	,371	1,000			
14-AÇİ4	,167	-,010	,049	-,044	,102	,036	-,022	-,063	-,117	-,037	,291	,475	,440	1,000		
15-AT5	,037	,002	,027	,011	,093	-,009	,344	,258	,188	,240	-,053	,019	,031	,095	1,000	
16-AT6	-,034	-,059	-,049	-,027	,029	-,043	,333	,180	,249	,208	,011	-,023	-,132	,019	,437	1,000

The highest correlation was observed between items AB2 and AB5 ($r = .568$). The lowest correlation was found between items AT1 and AB5 ($r = -.001$).

Confirmatory Factor Analysis (CFA)

The factors were analyzed through Confirmatory Factor Analysis (CFA) using the AMOS software, and the findings are presented below. The confirmatory factor analysis was conducted in three stages, and three different models were developed. The final path diagrams and modification indices of these models were calculated and are presented below.

Table 7. Model Fit Indices

Model No	X2	Sd	X2/Sd	p	CFI	GFI	RMR	SRMR	RMSEA	AGFI
1	237,604	101	2,353	,000	,866	,910	,056	,0608	,066	,878
2	200,084	99	2,021	,000	,901	,925	,042	,0566	,057	,896
3	186,733	98	1,905	,000	,913	,930	,041	,0551	,054	,902

High correlations between variables can lead to an increase in the chi-square (χ^2) value. The degrees of freedom (df) are critical in chi-square calculations, and the ratio of χ^2 to df is commonly used as a fit index. A ratio below 5 is generally considered indicative of a good model fit (Kelloway, 1998, pp. 23–40). According to Hu and Bentler (1999), a CFI value below .90 is considered acceptable. In the first model, the CFI was found to be .866. Schumacker and Lomax (2004) state that a GFI value above .90 indicates good fit; in the first model, this value was .910, as shown in the table.

According to Brown (2005) and Byrne (1994), an RMR value below .05 indicates excellent fit. The RMR value for the first model was .056, suggesting an acceptable fit. Similarly, a value below .05 for SRMR is regarded as excellent fit. The SRMR value in the first model was .0608, which again indicates acceptable fit.

As stated by Brown (2006), Jöreskog and Sörbom (1993), Raykov and Marcoulides (2008), Schumacker and Lomax (2004), and Sümer (2000), an RMSEA value below .05 indicates good fit, and values between .06 and .08 are considered acceptable. In the first model, the RMSEA was .066, falling within the acceptable range.

Hu and Bentler (1999) suggest that a CFI value above .90 represents good fit. In the second model, the CFI was .901. According to Schumacker and Lomax (2004), a GFI value above .90 also indicates good fit; in the second model, this value was .925. Brown (2005) and Byrne (1994) state that an RMR value below .05 indicates excellent fit. The second model had an RMR of .042, which suggests an acceptable fit.

Schermelleh-Engel et al. (2003) note that an SRMR value above .05 can still be considered indicative of acceptable fit; in the second model, the SRMR was .0566, supporting this interpretation. Furthermore, the RMSEA value in the second model was .057, which, according to Schermelleh-Engel et al. (2003), falls within the acceptable range of .05 to .08.

According to Hu and Bentler (1999), a CFI value above .90 indicates good model fit. In the third model, the CFI was calculated as .913. Schumacker and Lomax (2004) also state that a GFI value above .90 reflects good fit; in the third model, this value was .930, as shown in the table.

Brown (2005) and Byrne (1994) emphasize that an RMR value below .05 is indicative of excellent model fit. The RMR value for the third model was .041, which indicates an acceptable level of fit. According to Schermelleh-Engel et al. (2003), an RMR value above .05 can still reflect acceptable fit. The SRMR value in the third model was .0551, suggesting acceptable fit according to the same criteria.

Furthermore, Schermelleh-Engel et al. (2003) indicate that an RMSEA value between .05 and .08 is considered acceptable. In the third model, the RMSEA was found to be .054, supporting an acceptable level of model fit.

Table 8. Modification Index (MI) Values

Model	Items	Index	Percentage Change
1	e11 – e12	25.739	35.4%
	e9 – e10	10.753	30.5%
2	e3 – e4	12.665	10.0%

As a result of the modification index calculations, the MI value between e11 and e12 was found to be 25.739 with an expected parameter change of 35.4%, and between e9 and e10 the MI was 10.753 with a 30.5% change. Based on these results, the model was modified and the MI values were recalculated. In the second model, the MI value between e3 and e4 was 12.665, with a 10% expected change.

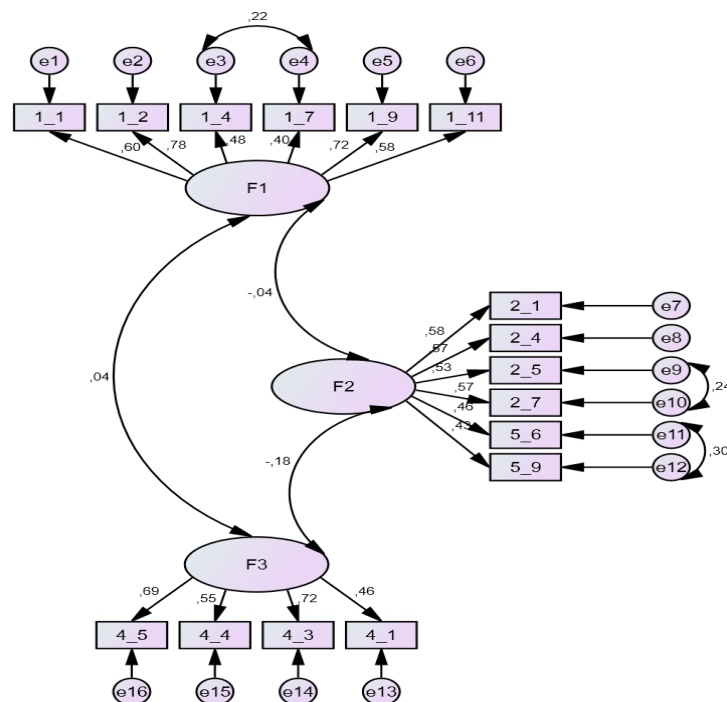


Figure 2. Path Diagram of the First Sample – CFA Model

Data Analysis

The data analyses presented in this section pertain to the Evaluation Scale of Family Education Activities in Preschool (OAEDÖ), which was developed through the revision of the Evaluation Scale of OBADER Family Education Sessions Delivered via Distance Education (UEOAEODÖ) created in the master's thesis study.

Table 9. Second Sample Group

Variables	Second Sample Group	N
Parent Gender	Female	208
	Male	33
Parent Age	20–25	5
	26–30	60
	31–35	123
	36–40	40
	41–45	13
	46–50	0
	51 and above	0
Parent Educational Level	Primary School	4
	Middle School	4
	High School	18
	Associate Degree	24
	Bachelor's Degree	158
	Postgraduate	33
Parent Occupation	Not Employed	70
	Public Servant	94
	Worker	48
	Tradesperson	29

Among the 241 parents in the second sample group, 208 were female and 33 were male. The highest number of participants (123) were in the 31–35 age group. In terms of educational background, the majority held a bachelor's degree (158 participants). Regarding occupational status, the largest group consisted of public servants (94 participants).

Table 10. EFA Results for the Second Sample

KMO and Bartlett's Test		
Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		,747
Bartlett's Test of Sphericity	Approx. Chi-Square	1042,404
	df	120
	Sig.	,000

The KMO value was found to be 0.747, indicating a good level of sampling adequacy. This suggests that the dataset possesses a sufficient sample size for conducting factor analysis.

Table 11. Communalities

Communalities		
	Initial	Extraction
AB1	1,000	,708
AB2	1,000	,656
AB3	1,000	,439
AB4	1,000	,359
AB5	1,000	,591
AB6	1,000	,584
AT1	1,000	,471
AT2	1,000	,425
AT3	1,000	,358
AT4	1,000	,418
AT5	1,000	,491
AT6	1,000	,396
AÇİ1	1,000	,379
AÇİ2	1,000	,620
AÇİ3	1,000	,432
AÇİ4	1,000	,633

Following the factor analysis, an examination of the extraction values revealed that the lowest communalities were observed for AÇİ1 (.379), AB4 (.359), and AT3 (.359). These values indicate the extent to which each variable is explained by the existing factor structure.

Table 12. Total Variance Explained

Component	Initial Eigenvalues			Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	3,312	20,702	20,702	3,312	20,702	20,702	3,303	20,644	20,644
2	2,562	16,014	36,716	2,562	16,014	36,716	2,491	15,567	36,211
3	2,085	13,031	49,747	2,085	13,031	49,747	2,166	13,536	49,747
4	1,114	6,964	56,710						
5	,968	6,049	62,760						
6	,872	5,452	68,212						
.....									
.....									

This table summarizes the impact of the extracted factors on variance, as well as the percentage of variance explained by each factor. Factor 1, with an eigenvalue of 3.312, accounted for 20.702% of the total variance. Factor 2, with an eigenvalue of 2.562, explained 16.014% of the total variance. Factor 3, with an eigenvalue of 2.085, accounted for 13.031% of the total variance. These results indicate that a significant portion of the variance among variables is explained by the three-factor structure.

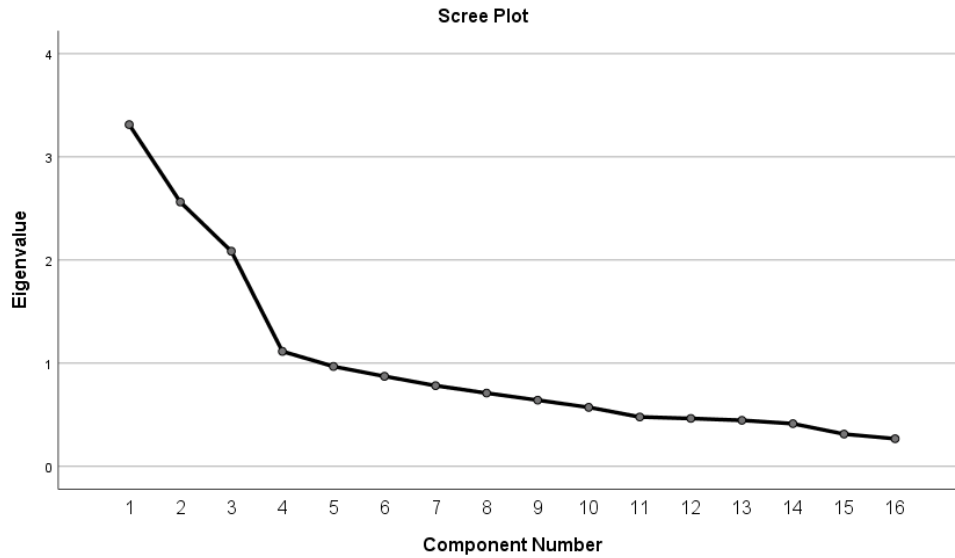


Figure 3. Scree Plot of the Second Sample

An examination of the scree plot suggests that a three-factor solution is appropriate.

Table 13. Rotated Component Matrix

	Rotated Component Matrix		
	Component		
	1	2	3
AB1	,838	-,049	-,063
AB2	,801	-,103	-,058
AB5	,754	,050	,139
AB6	,752	-,026	,133
AB3	,660	,059	-,020
AB4	,595	,016	-,065
AT5	,021	,687	,136
AT1	,008	,687	,002
AT2	-,068	,630	-,155
AT4	,128	,624	-,108
AT6	-,076	,621	,065
AT3	-,025	,582	-,136
AÇİ4	,007	,035	,795
AÇİ2	-,052	,001	,786
AÇİ3	,000	-,104	,649
AÇİ1	,060	-,072	,608

Table 13 presents the strength of the relationships between each variable and the three extracted factors. Loading values close to 1 indicate a strong relationship between the variable and the respective factor. Negative loadings suggest a negative relationship with that factor. As seen in the table, the loadings for the first factor range between .838 and .595, for the second factor between .687 and .582, and for the third factor between .795 and .608.

Based on the reliability analysis, the Cronbach's Alpha coefficient was found to be .778 for the *parental expectations* factor, .710 for the *parental attitudes* factor, and .652 for the *parent-child communication* factor. The overall reliability coefficient for the scale was calculated as .624.

Table 14. Model Fit Indices for the Second Sample

Model No	X2	Sd	X2/Sd	p	CFI	GFI	RMR	SRMR	RMSEA	AGFI
1	218,051	101	2,159	,000	,877	,877	,061	,0635	,069	,861
2	193,966	100	1,940	,000	,901	,910	,049	,0599	,063	,877
3	187,870	99	1,898	,000	,906	,913	,046	,0592	,061	,880

High correlations between variables can lead to an increase in the chi-square (χ^2) value. Degrees of freedom (df) are a critical parameter in chi-square calculations, and the ratio of χ^2 to df is commonly used as a model fit index. In the literature, a χ^2 /df ratio below 5 is generally considered indicative of acceptable model fit.

A CFI value below .90 is regarded as acceptable in the literature. In the first model, the CFI was found to be .877. A GFI value above .90 is considered a sign of good fit. In the first model, as shown in **Table 14**, the GFI was also .877.

According to the literature, an RMR value below .05 is indicative of excellent model fit. In the first model, the RMR was found to be .061. An SRMR value below .05 is also considered indicative of excellent fit. The SRMR value of .0635 in the first model suggests an acceptable fit. RMSEA values below .05 indicate good fit, while values between .06 and .08 indicate acceptable fit. The RMSEA in the first model was .069, which falls within the acceptable range.

In the second model, the CFI was found to be .901, indicating good model fit. The GFI was calculated as .910, as shown in Table 14. The RMR value of .049 and SRMR value of .0599 in the second model also indicate acceptable fit. The RMSEA value in the second model was .063, again reflecting an acceptable level of fit.

In the third model, the CFI value was .906, and the GFI was .913, as indicated in Table 14. The RMR value of .046 and SRMR value of .0592 both suggest acceptable fit. The RMSEA value in the third model was .061, which falls within the acceptable range of .05 to .08, indicating a satisfactory model fit.

Model	Items	Index	Percentage Change
1	e7 – e8	17.310	30.8%
2	e8 – e10	4.803	-20.1%

As a result of the modification index calculations, the MI value between e7 and e8 was found to be 17.310 with an expected parameter change of 30.8%. The MI value between e8 and e10 was 4.803, with a parameter change of -20.1%.

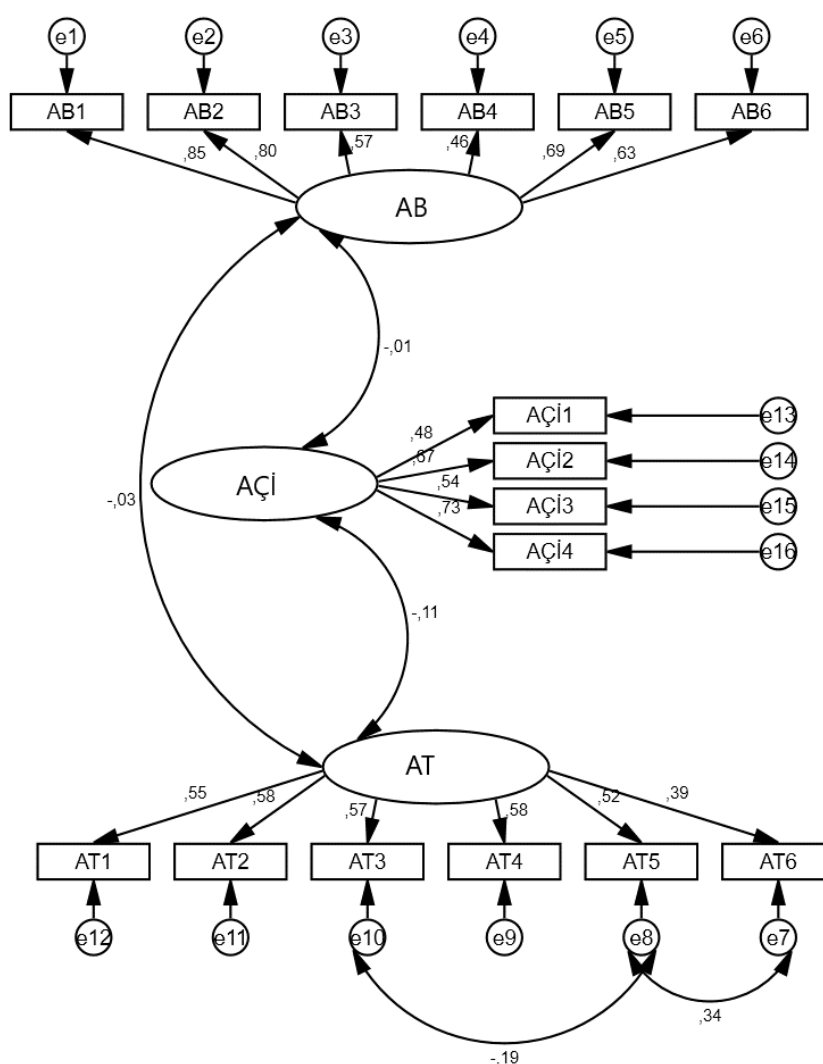


Figure 4. Path Diagram of the Second Sample – CFA Model

RESULTS, DISCUSSION, AND RECOMMENDATIONS

It can be stated that parental expectations for preschool education, parental attitudes, and parent-child communication significantly influence children's social and cultural development (Özyürek, 2015). In this context, gathering feedback from families plays a crucial role in understanding and meeting the needs of both parents and children. This, in turn, contributes to the improvement of educational practices (Brown et al., 2016).

Discussion of Findings Related to the *Parental Expectations* Sub-Dimension of the Scale

Within this framework, it can be suggested that families consider it important for preschool education to address their children's developmental characteristics and needs, to support proper use of the Turkish language, to prepare children for primary school, to maintain their learning motivation, and to keep parents regularly informed about their children's development.

The literature review indicates that parental expectations regarding preschool education often include support for developmental domains, the presence of qualified teachers, preparation for primary education, high-quality physical conditions of institutions, and regular communication from schools about children's educational progress (Gülender, 1993; Aktaş Arnas, 2002; Hinnant et al., 2009; Erşan, 2019; Rubie-Davies et al., 2006; Sabırlı Özışıklı, 2008; Buldu & Güner Pekacar, 2023; Sevinç, 2006; Seyfullahoğulları, 2012).

Moreover, expectations concerning the development of cognitive, self-care, and social skills necessary for primary school readiness are also emphasized (Metin et al., 1993; Şimşek & İvrendi, 2014). In a study conducted by Yalman (2024), it was reported that more than half of the mothers expressed a desire to receive information about child education.

Yankayış and Yankayış (2024) found that parental expectations from preschool education also included contributing to children's socialization, fostering cooperation, sharing, and responsibility, and strengthening commitment to national and moral values. In addition, expectations regarding the development of children's language skills were noted. According to the findings, parents expressed their desire for their children to become individuals who speak Turkish accurately and fluently.

When the current study is examined, it is evident that numerous studies in the literature focus specifically on the factor of parental expectations. As reflected in the paragraphs above, the items under the parental expectations factor in this study are consistent with those emphasized in previous research.

Discussion of Findings Related to the *Parental Attitudes* Sub-Dimension of the Scale

The literature indicates various types of parental attitudes, including democratic, perfectionist, overly permissive, authoritarian, neglectful, inconsistent, and overprotective parenting styles. Below is a discussion of relevant findings from the literature on these classifications.

One of the fundamental factors influencing the relationship between parents and children is the attitudes and behaviors of parents. Through these behaviors, parents serve as role models for their children (Alrehaly, 2011; Zöhrap, 2004). In order for children to exhibit consistent behavior, become independent individuals, be self-sufficient, and form healthy relationships, they must have healthy interactions with their parents (MoNE, 2013b). Parental attitudes and the quality of relationships within their social environment affect the development of both positive and negative behaviors in children (Özyürek, 2004; Arslan, 2022; Çağdaş, 2015; Attili et al., 2011; Yılmaz Bolat, 2017). All of these aspects are directly related to parenting attitudes and intra-family relationships.

In the democratic parenting style, parents are attentive and warm toward their children, and they include their children in the decision-making processes within the family. Authoritarian parenting is characterized by excessive control and rigid rules, with limited emotional warmth in interactions. Neglectful parenting involves a lack of boundaries and indifference to the child's needs and demands. Overly permissive parents fulfill their children's requests even when they are unreasonable, prioritizing the child's desires over their own. In inconsistent parenting, one parent may deem a particular behavior acceptable while the other does not, leading to conflicting signals. Overprotective parents excessively shelter and control their children, often performing tasks that the child could do independently, thus hindering experiential learning. In the perfectionist style, parents have high expectations of their children, which may cause the child to feel inadequate and hinder the development of a healthy self-concept (Baumrind, 1966; MoNE, 2013b; Demircioğlu & Ömeroğlu, 2014).

A review of the studies mentioned below shows that the impact of parental attitudes on children is undeniable.

In a study by Tönbül (2019), it was found that participation in a parental education program helped reduce strict disciplinary practices among mothers. Demircioğlu and Ömeroğlu (2014) observed that both educational parenting training and peer-to-peer training approaches had a positive impact on family functioning and parenting styles in their experimental group. Turan et al. (2019) conducted a study with a mother of a child with hearing loss and found that after the parent education program, there was an increase in the mother's positive behaviors supporting the child's development, as well as favorable changes in the educator's practices. The results of these studies support the conclusion that the parental attitudes factor, developed as part of the present study, plays a significant role in shaping a child's future development.

Discussion of Findings Related to the Parent-Child Communication Sub-Dimension of the Scale

A review of the literature reveals that parent-child communication emphasizes key elements such as establishing eye contact, understanding emotions, using age-appropriate and developmentally appropriate language, and the importance of bidirectional communication. Relevant findings from the literature are discussed below.

During the preschool period, it is critical for parents to serve as role models for their children. Parents who act as role models also play a crucial role in fostering effective communication skills in their children. The importance of effective communication between parents and children is widely recognized—not only by professionals and educators in the field but also by society at large (Erkan, 2020; Çağdaş & Şahin Seçer, 2015). Children who acquire effective communication skills are better able to express their emotions and thoughts clearly, use language appropriately, and interact successfully with others.

There are numerous studies in the literature focused on family communication (Baxter & Akkoor, 2011; Arabacı & Ömeroğlu, 2016; Arabacı et al., 2022; Büyükmedar, 2022; Çınar, 2022; Özler, 2023; Öztürk, 2023; Önder et al., 2015). These studies highlight the necessity of communication skills for individuals to build healthy relationships. Key components of such communication include active listening, effective speaking, showing empathy, being open to feedback, and using “I-language” and a language of acceptance rather than accusatory or aggressive speech (Güven, 2016). As demonstrated in these studies, effective communication skills contribute to overcoming or reducing communication barriers (Şahin & Aral, 2012; Doğan & Sinan, 2024).

The findings support the importance of the parent-child communication dimension developed in this scale. This dimension captures a range of communication-related behaviors and expectations grounded in the literature.

Conclusion and Recommendations

In conclusion, the scale developed through this study integrates various dimensions that have previously been addressed individually in the literature. From this perspective, the scale is significant in that it addresses three core topics in family education—parental expectations, parental attitudes, and parent-child communication—in a unified structure.

The scale developed as part of this research can be tested by other researchers using different sample groups. Moreover, in future studies, the dimensions of parental expectations, parental attitudes, and communication with children can be compared or used as analytical tools in broader educational contexts.

Ethical approval for conducting this study was obtained from the Kastamonu University Social and Human Sciences Research and Publication Ethics Committee with the decision number 56, dated 25.03.2021.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, author-ship, and/or publication of this article.

Statements of publication ethics

We hereby declare that the study has not unethical issues and that research and publication ethics have been observed carefully.

Researchers' contribution rate

The study was conducted and reported with equal collaboration of the researchers.

Ethics Committee Approval Information

For this study, ethics committee permission was obtained from Kastamonu University Social and Human Sciences Research and Publication Ethics Committee numbered 1/56 and dated 25.03.2021.

REFERENCES

- Aktaş Arnas, Y. (2002). Velilerin okul öncesi eğitime ve okul öncesi eğitim kurumlarına yönelik tutumları. *Çukurova Üniversitesi Sosyal Bilimler Enstitüsü Dergisi*, 10(10), 24-36.
- Alrehaly, E. D. (2011). *Parental attitudes and the effects of ethnicity: how they influence children's attitudes toward science education*. Arkansas: The University of Arkansas.
- Arabacı, N. & Ömeroğlu, E. (2016). Anne-Baba-Çocuk İletişimi Değerlendirme Aracı'nın (ABÇİDA) geliştirilmesi: Geçerlik güvenirlik çalışması. *Uluslararası Türk Eğitim Bilimleri Dergisi*, 4(7), 1-21.
- Arabacı, N., Özyürek, A. & Gözün Kahraman, Ö. (2022). Okul öncesi dönem çocuklarının mizaç özellikleri ve baba-çocuk iletişimi arasındaki ilişkinin incelenmesi. *Nevşehir Hacı Bektaş Veli Üniversitesi SBE Dergisi*, 12(2), 656-675.
- Arslan, A. (2022). Erken çocukluk döneminde çocuğun gelişiminin desteklenmesinde aile. *Gümüşhane Üniversitesi Sağlık Bilimleri Dergisi*, 11(4), 1676-1684.

- Attili, G., Vermigli, P. & Roazzi, A. (2011). Rearing styles, parents' attachment mental state, and children's social abilities: The link to peer acceptance. *Child Development Research*, <https://doi.org/10.1155/2011/267186>.
- Baumrind, D. (1966). Effects of authoritative parental control on child behavior. *Child Development*, 37(4), 887-907.
- Baxter, L. A. & Akkoor, C. (2011). Topic expansiveness and family communication patterns. *Journal of Family Communication*, 11(1), 1-20.
- Böyükmedar, P. (2022). *İletişim becerileri eğitim programının baba-çocuk ilişkisi üzerine etkisinin incelenmesi*. [Yayınlanmamış Yüksek Lisans Tezi]. Çanakkale Onsekiz Mart Üniversitesi.
- Brown, E. L., Mahatmya, D. & Vesely, C. K. (2016). Home and school influences on the behavioral and academic outcomes of low-income children of color. *Journal of Children and Poverty*, 22(2), 93-112.
- Brown, J. D. (2005). Statistics Corner. Questions and answers about language testing statistics: Characteristics of sound qualitative research. *Shiken: JALT Testing & Evaluation SIG Newsletter*, 9(2), 31-32. Also retrieved from the World Wide Web at http://www.jalt.org/test/bro_22.htm
- Brown, J. D. (2006). Statistics Corner. Questions and answers about language testing statistics: Generalizability from second language research samples. *Shiken: JALT Testing & Evaluation SIG Newsletter*, 10(2), 21-22. Also retrieved from the World Wide Web at http://www.jalt.org/test/bro_24.htm
- Buldu, M. & Güner-Pekacar, R. (2023). İdeal bir okul öncesi eğitim ortamı: Ebeveynlerin okul öncesi eğitim kurumlarından beklentileri. *Van Yüzüncü Yıl Üniversitesi Eğitim Fakültesi Dergisi*, 20(2), 625-643.
- Büyüköztürk, Ş., Kılıç Çakmak, E., Akgün, Ö. E., Karadeniz, Ş. & Demirel, F. (2020). *Bilimsel araştırma yöntemleri* (28. Baskı). Pegem Akademi.
- Byrne, B. M. (1994). Burnout: Testing for the validity, replication, and invariance of causal structure across elementary, intermediate, and secondary teachers. *American Educational Research Journal*, 31(3), 645-673.
- Can, A. (2013). *SPSS ile bilimsel araştırma sürecinde nicel veri analizi*. Pegem Akademi.
- Çağdaş, A. (2015). *Anne-baba-çocuk iletişimi*. Eğiten Kitap.
- Çağdaş, A. & Şahin Seçer, Z. (2015). *Anne baba eğitimi*. Eğiten Kitap Yayıncılık.
- Çınar, E. (2022). *Okul öncesi eğitim kurumlarına devam eden 48-72 aylık çocukların anne-baba-çocuk iletişim becerileri ile sosyal problem çözme becerileri arasındaki ilişkinin bazı değişkenlere göre incelenmesi*. [Yayınlanmamış Yüksek Lisans Tezi]. Gazi Üniversitesi.
- Demircioğlu, H. & Ömeroğlu, E. (2014). Eğitici anne eğitimi ve anneden anneye eğitim yaklaşımlarının annelerin aile işlevleri ve çocuk yetiştirme tutumları üzerindeki etkileri. *Kuram ve Uygulamada Eğitim Bilimleri Dergisi*, 14(4), 1439-1465.
- Doğan, F. N. & Sinan, A. T. (2024). Etkili iletişim becerileri ölçeği: Geçerlik ve güvenilirlik çalışması. *International Journal of Eurasia Social Sciences*, 15(57), 1071-1087.
- Erkan, S. (2020). Aile ve aile eğitimi ile ilgili temel kavramlar. Z. F. Temel (Ed.). *Aile eğitimi ve erken çocukluk eğitiminde aile katılım çalışmaları* (s. 1-34) içinde. Anı Yayıncılık.
- Erşan, Ş. (2019). Ebeveynlerin okul öncesi eğitim kurumundan ve öğretmenlerden beklentilerinin incelenmesi. *Sinop Üniversitesi Sosyal Bilimler Dergisi*, 3 (2), 161-178.
- Gülender, S. (1993). *Vellilerin okul öncesi eğitim kurumlarından beklentileri*. 9. Ya-Pa Okul Öncesi Eğitimi ve Yaygınlaştırılması Semineri. Ya-Pa Yayıncılık.
- Güven, A. Z. (2016). Etkili iletişimin kuralları ve etkili iletişim kurma süreci. B. Güven (Ed.). *Etkili İletişim* (s. 53-66) içinde. Pegem Akademi.
- Hinnant, J. B., O'Brien, M. & Ghazarian, S. R. (2009). The longitudinal relations of teacher expectations to achievement in the early school years. *Journal of Educational Psychology*, 101(3), 662-670.
- Hinton, P. R., Brownlow, C., McMurray, I. & Cozens, B. (2004). *SPSS Explained*. (First Edition). Routledge.
- 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.
- Jöreskog, K. G. & Sörbom, D. (1993). *LISREL 8: structural equation modeling with the SIMPLIS command language*. Scientific Software International.
- Kelloway, E. K. (1998). *Assessing model fit. Using Lisrel for structural equation modeling*. (Third Edition). Sage Publications.
- Koçak, A. & Sallabaş, M. E. (2022). Türkçe öğretmen adaylarının uzaktan eğitim uygulamalarına yönelik tutum ölçeğinin geliştirilmesi geçerlilik ve güvenilirlik çalışması. *Bandırma Onyedli Eylül Üniversitesi Sosyal Bilimler Araştırmaları Dergisi*, 5(2), 125-139.
- Mahoney, G., Kaiser, A., Girolametto, L., MacDonald, J., Robinson, C., Safford, P. & Spiker, D. (1999). Parent education in early intervention: A call for a renewed focus. *Topics in Early Childhood Special Education*, 19(3), 131-140.
- Metin, N., Arı, M., Güneysu, S., Dikmen, B., Atik, B., Aydın, C., Üstün, E. & Uysal, N. (1993). *Anne babaların anaokulundan beklentileri*. 9. Ya-Pa Okul Öncesi Eğitimi ve Yaygınlaştırılması Semineri. Ya-Pa Yayıncılık.
- Milli Eğitim Bakanlığı Temel Eğitim Genel Müdürlüğü (2002). *Okul öncesi eğitim programı*. Ankara.
- Milli Eğitim Bakanlığı Temel Eğitim Genel Müdürlüğü (2006). *Okul öncesi eğitim programı*. Ankara.
- Milli Eğitim Bakanlığı (2013a). *Okul öncesi eğitim programı*. Ankara.
- Milli Eğitim Bakanlığı (MEB), (2013b). *Okul öncesi eğitim programı ile bütünleştirilmiş aile destek eğitim rehberi (OBADER)*. T.C. Millî Eğitim Bakanlığı, Temel Eğitim Genel Müdürlüğü.
- Milli Eğitim Bakanlığı (MEB), (2013c). *0-36 ay çocuklar için eğitim programı ile bütünleştirilmiş aile destek eğitim rehberi (EBADER)*. Ankara.
- Milli Eğitim Bakanlığı (2024a). *Okul öncesi eğitim programı*. Ankara.
- Milli Eğitim Bakanlığı (2024b). *Türkiye yüzümlü maarif modeli okul öncesi eğitim programı*. Ankara.
- Nunnally, J. C. & Bernstein, I. H. (1994). *Psychometric theory* (3rd ed.). New York, NY: McGRAW-HILL.
- Önder, A., Balaban Dağal, A. & Şallı, D. (2015). 5-6 yaş çocukları için iletişim becerileri ölçeği geçerlik-güvenirlik çalışması. *Eğitim ve Öğretim Araştırmaları Dergisi*, 4(1), 1-13.
- Özler, E. (2023). *60-72 aylık çocukların benlik algıları ile anne-baba-çocuk iletişimi arasındaki ilişkinin incelenmesi*. [Yayınlanmamış Yüksek Lisans Tezi]. Mersin Üniversitesi.
- Öztürk, E. (2023). *Aile içi iletişimin çocuk üzerindeki etkileri*. [Yayınlanmamış Yüksek Lisans Tezi]. KTO Karatay Üniversitesi.
- Özyürek, A. (2004). *Kırsal bölge ve şehir merkezinde yaşayan 5-6 yaş grubu çocuğa sahip anne-babaların çocuk yetiştirme tutumlarının incelenmesi*. [Yayınlanmamış Yüksek Lisans Tezi]. Gazi Üniversitesi.

- Özyürek, A. (2015). Okul öncesi çocukların sosyal beceri düzeyleri ile anne tutumları arasındaki ilişkinin incelenmesi. *Milli Eğitim Dergisi*, 45(206), 106-120.
- Özyürek, A., Adıbatmaz, M., Çetin, A., Meriç, A., Kılınç, N., Erzurumluoğlu, Ş., Yücesan, Y. & Atalay, D. (2015). Okul öncesi eğitim kurumu ve ilkokullarda aile eğitimi ve katılımı çalışmalarının incelenmesi. *The Journal of Social Sciences*, 2(3), 28-43.
- Raykov, T. & Marcoulides, G. A. (2008). An introduction to applied multivariate analysis. (First Edition). Taylor & Francis Group.
- Rubie-Davies, C., Hattie, J. & Hamilton, R. (2006). Expecting the best for students: Teacher expectations and academic outcomes. *British Journal of Educational Psychology*, 76(3), 429-444.
- Sabırlı Özışıklı, I. (2008). A study of parent involvement in the Boğaziçi University preschool center. [Master's thesis]. Boğaziçi University.
- Schermelleh Engel, K., Moosbrugger, H. & Müller, H. (2003). Evaluating the fit of structural equation models: Tests of significance and descriptive goodness-of-fit measures. *Psychological Research Online*, 2 (8), 23-74.
- Schumacker, R. E. & Lomax, R. G. (2004). A Beginner's Guide to Structural Equation Modeling. (Second Edition). Lawrence Erlbaum Associates.
- Sevinç, M. (2006). Okul öncesi eğitimi alan çocukların annelerinin okuldan beklentileri. *Atatürk Üniversitesi Kazım Karabekir Eğitim Fakültesi Dergisi*, 13, 218-225.
- Seyfullahoğlu, A. (2012). Ailelerin anaokullarından beklentileri üzerine bir araştırma. *Marmara Sosyal Araştırmalar Dergisi*, (2), 1-15.
- Slotkin, R., Bierman, K.L., Heinrichs, B.S. et al (2024). Parent Outreach Efforts Extend the Sustained Benefits of a Preschool Classroom Intervention: Adolescent Follow-Up of a Randomized Controlled Trial. *Prev Sci* 25, 1217–1227. <https://doi.org/10.1007/s11121-024-01758-8>
- Smith, C., Perou, R., ve Lesene, C. (2002). Parent education. Bornstein, M. H (Eds.), *Handbook of Parenting Volume 4 Social Conditions and Applied Parenting*, (pp. 389-409). New Jersey: Lawrence Erlbaum Associates, Inc., Publishers.
- Sümer, N. (2000). Yapısal Eşitlik Modelleri: Temel Kavramlar ve Örnek Uygulamalar (Structural Equation Modeling: Basic Concept and Applications). *Türk Psikoloji Yazıları*, 3(6), 49-74.
- Şahin, S. & Aral, N. (2012). Aile içi iletişim. *Ankara Sağlık Bilimleri Dergisi*, 1(3), 55-66.
- Şimşek, Z. C. & İvrendi, A. (2014). Ebeveynlerin okul öncesi eğitim kurumlarından beklentileri. *Hacettepe Üniversitesi Eğitim Fakültesi Dergisi*, 29(2), 240-254.
- Tezel Şahin, F. & Özyürek, A. (2008). *Anne baba eğitimi ve okul öncesinde aile katılımı*. Morpa Kültür Yayınları.
- Tönbül, Ö. (2019). Aile eğitimi programının annelerin çocuk yetiştirme tutumlarına etkisinin incelenmesi. *Okul Psikolojik Danışmanlığı Dergisi*, 2(1), 46-72.
- Turan, Z., Koca, A. & Uzun, Y. (2019). İşitme kayıplı çocuğu olan bir annenin aile eğitimi sürecinin incelenmesi. *Ankara Üniversitesi Eğitim Bilimleri Fakültesi Özel Eğitim Dergisi*, 20(1), 93-117.
- Uşun, S. (2006). *Uzaktan eğitim*. Nobel Yayın Dağıtım.
- Ünal, F. (2018). Aile eğitimi. (A. B. Aksoy, Ed.). *Aile eğitimi ve katılımı* içinde (s. 12-40). Hedef CS Basın Yayın.
- Ünlü Çetin, Ş. (2016). Okul öncesi eğitim programı ile bütünleştirilmiş aile destek eğitim rehberi'nde (OBADER) toplumsal cinsiyet eşitliği vurgusunun baba katılımı özelinde incelenmesi. *Current Research in Education*, 2(2), 61-83.
- Yalman, D. (2024). *MEB aile eğitim programına katılan ve katılmayan okul öncesi eğitim çağında çocuğu olan annelerin aile eğitim programlarından beklentileri ve görüşleri*. [Doktora tezi]. Marmara Üniversitesi.
- Yankayış, S. & Yankayış, Z. (2024). Okul öncesi eğitim kurumlarına devam eden çocukların ailelerinin kurumlara yönelik beklentileri. *Uluslararası Öğrenen Toplum Dergisi*, 1(1), 52-70.
- Yılmaz Bolat, E. (2017). *Okul öncesi eğitimde aile eğitimi ve katılımı*. Eğiten Kitap.
- Zöhrap, D. U (2004). *Lise öğrencilerinde algılanan anababa tutumları ve aile yapısının ergen kişilik özelliklerine etkisi*. [Yüksek lisans tezi]. Marmara Üniversitesi.



| Research Article / Araştırma Makalesi |

Examination of Writing Skill-Focused Learning Outcomes in Turkish and Social Studies Curricula in Terms of Cognitive Process Steps

Türkçe ve Sosyal Bilgiler Öğretim Programlarındaki Yazma Becerisi Odaklı Öğrenme Çıktılarının Bilişsel Süreç Basamakları Açısından İncelenmesi

Tahsin Yıldırım¹, Mehmet Fidan²

Keywords

1. Turkish language course curriculum
2. Social studies course curriculum
3. Cognitive strategies
4. Writing skills

Anahtar Kelimeler

1. Türkçe dersi öğretim programı
2. Sosyal bilgiler dersi öğretim programı
3. Bilişsel stratejiler
4. Yazma becerisi

Received/Başvuru Tarihi
18.03.2025

Accepted / Kabul Tarihi
10.10.2025

Abstract

Purpose: The main objective of this research is to examine the writing skill-focused learning outcomes in the middle school Turkish language curriculum and social studies curriculum, prepared within the framework of the Turkey Century Education Model, in terms of cognitive process steps.

Design/Methodology/Approach: The research was conducted using the document analysis method, one of the qualitative research methods. The data sources of the research consist of the Turkish language curriculum and social studies curriculum. The curricula were analyzed using the content analysis technique in terms of learning outcomes related to writing skills. In determining the learning outcomes for writing skills, those directly related to the writing skill area in the Turkish language course were interpreted within the scope of cognitive domain steps. In the field of social studies, explanations for learning outcomes were examined to determine whether these explanations were related to the writing skill area. Following these determinations, learning outcomes related to the writing skill area within the scope of the social studies course were examined in terms of cognitive domain steps.

Findings: According to the findings obtained from the examination of the curricula, it was determined that the process components of the learning outcomes in the Turkish language curriculum, while having a close distribution in terms of lower-level and higher-level cognitive domain steps, were mostly oriented towards the application step and higher-level cognitive domain steps such as analyze and evaluate. In the learning outcomes related to the writing area in the social studies curriculum, it was determined that the understand step was significantly prominent compared to other skill areas.

Highlights: It is thought that the balanced distribution of learning outcomes related to the writing skill area in terms of cognitive domain dimensions is important for acquiring the skill. It is believed that reflecting this situation in the curricula will have a positive effect on students' written expression skills.

Öz

Çalışmanın amacı: Bu araştırmanın temel amacı Türkiye Yüzyılı Maarif Modeli çerçevesinde hazırlanan ortaokul Türkçe dersi öğretim programında ve sosyal bilgiler dersi öğretim programında yer alan yazma becerisi odaklı öğrenme çıktılarının bilişsel süreç basamakları açısından incelenmesidir.

Materyal ve Yöntem: Araştırma nitel araştırma yöntemlerinden doküman incelemesi yöntemiyle gerçekleştirilmiştir. Araştırmanın veri kaynaklarını Türkçe dersi öğretim programı ve sosyal bilgiler dersi öğretim programları oluşturmaktadır. Öğretim programları yazma becerisine yönelik öğrenme çıktıları açısından içerik analizi tekniği ile incelenmiştir. Yazma becerisine yönelik öğrenme çıktılarının belirlenmesinde Türkçe dersi alanında doğrudan yazma becerisi alanına yönelik öğrenme çıktıları bilişsel alan basamakları kapsamında yorumlanmıştır. Sosyal bilgiler dersi alanında ise öğrenme çıktılarına yönelik açıklamalar incelenerek bu açıklamaların yazma becerisi alanına yönelik olup olmadığı belirlenmiştir. Bu tespitlerin ardından sosyal bilgiler dersi kapsamında yazma becerisi alanına yönelik öğrenme çıktıları bilişsel alan basamakları açısından incelenmiştir.

Bulgular: Öğretim programlarının incelenmesi sonucu elde edilen bulgulara göre Türkçe dersi öğretim programındaki öğrenme çıktılarının süreç bileşenlerinin, alt düzey ve üst düzey bilişsel alan basamakları açısından yakın dağılım olmakla birlikte çoğunlukla uygulama basamağına ve analiz et, değerlendir gibi üst düzey bilişsel alan basamaklarına yönelik olduğu tespit edilmiştir. Sosyal bilgiler dersi öğretim programındaki yazma alanına yönelik öğrenme çıktılarında anla başağının diğer beceri alanlarına göre belirgin bir biçimde öne çıktığı belirlenmiştir.

Önemli Vurgular: Yazma becerisi alanına yönelik öğrenme çıktılarının bilişsel alan boyutları açısından dengeli dağılımının becerinin kazandırılması açısından önemli olduğu düşünülmektedir. Bu durumun öğretim programlarına yansıtılmasının öğrencilerin yazılı ifade becerilerine olumlu etkisinin olacağı düşünülmektedir.

¹ Corresponded Author, Aksaray University, Faculty of Education, Department of Social Studie Education, Aksaray, Türkiye; <https://orcid.org/0000-0002-4130-4874>

² Aksaray University, Faculty of Education, Department of Turkish Education, Aksaray, Türkiye; <https://orcid.org/0000-0002-9347-2992>

INTRODUCTION

Language education, which begins with the family, aims to develop students' ability to express themselves correctly and understand what is expressed accurately through specific plans and programs during the school period. Indeed, this situation is also included in the Turkish Qualifications Framework, which is one of the sources on which the curriculum is based, and this skill area is addressed within the framework of "Communication in the mother tongue" competence (Ministry of National Education [MoNE], 2019). When examining the competence of communication in the mother tongue, emphasis is placed on individuals' ability to communicate correctly with their environment both in educational processes and in daily and social life (MoNE, 2019). An important expression skill in this competence area is writing (MoNE, 2019). The development of writing skills, which stands out as an important skill area in self-expression, requires diversity in terms of methods and techniques, as in other language skill areas (MoNE, 2019). When examining writing methods and techniques, there are various activities that can be applied during the course process, such as note-taking, summarizing, free writing, creative writing, writing based on senses, and critical writing (MoNE, 2006). In addition to these methods, various activities such as text completion, rewriting by changing the narrator of the text, rewriting by changing the basic elements of the text, and reconstructing the text in a different genre can be applied specific to the topic and learning outcome covered in the lesson (Gündüz & Şimşek, 2016).

It is observed that writing a good text is associated with cognitive development (Kellogg, 2008). The writing process, which has a holistic structure, is an area that requires the use of cognitive skills along with various stages such as planning and revision (Klimova, 2014). This situation ensures that writing skills maintain an important position throughout students' educational processes. Indeed, there appear to be various studies aimed at developing students' writing skills. Kellogg and Raulerson (2007) draw attention to deliberate practice in text writing, indicating that this may have an effect on students gaining cognitive control over the text. It is seen that the development of cognitive skills also affects students' writing abilities. In his research, Rao (2007) states that brainstorming strategy instruction contributes to the improvement of students' writing skills. There are also studies on how metacognitive strategies improve students' writing skills (Cer, 2019).

In the field of writing, students not only develop their ability to convey information but also enhance their competence in creating texts, and accordingly, their aesthetic capabilities in the writing domain also improve (Ungan, 2007). In writing education, text creation is viewed as a structured process with various stages, requiring operations such as text planning, text generation, and text review (Kellogg, 2008). Kellogg and Raulerson (2007) emphasize the use of deliberate practice in developing students' writing skills, noting that such practices contribute to enhancing cognitive control in the text production process. The writing education process encompasses various principles from the initial stage to text creation. Primarily, one of the issues teachers should pay attention to in the text creation process is to positively develop students' motivation towards writing. Gündüz and Şimşek (2016) state that activities that support self-confidence and motivate should be included to develop students' motivation towards writing. Additionally, principles such as developing the ability to effectively use one's native language, acquiring a unique style, being a good observer, gaining regular reading habits, and questioning and internalizing knowledge are also emphasized in developing writing skills (Gündüz & Şimşek, 2016). In writing education, besides positively developing students' motivation towards the process, attention should also be paid to considering individual differences and ensuring students' active participation in activities (MoNE, 2019). When examining studies related to this research's subject area, Karagöl and Tarakcı (2018) address various issues regarding writing outcomes in Turkish textbooks. Karacaoğlu, Dağ, and Uzun (2021) found that writing activities in the Turkish textbooks they examined were more product-oriented. Coşkun and Nariñç (2018) state that writing activities are given more space compared to speaking activities in the Turkish textbook.

It is important to identify the cognitive domain in which the competencies expected from students are located. Indeed, in this area, which is included in the curriculum as cognitive strategies and metacognitive strategies, emphasis is placed on students possessing these skills and being able to use them independently (MoNE, 2024b). In this context, the main purpose of the research is to determine the writing skill-focused learning outcomes, i.e., the basic competencies expected from students, with a focus on cognitive strategies in the middle school Turkish language curriculum (MoNE, 2024b) and Social Studies curriculum (MoNE, 2024a) prepared within the framework of the Turkish Century Education Model. In this context, it is observed that the Middle School Turkish Language Curriculum includes the cognitive strategies for writing skills under the headings "Planning, Text Creation, Revision, Sharing" (MoNE, 2024b, p. 10). Additionally, learning outcomes related to writing skills in the field of social studies, another subject where written expression skills are frequently preferred, will be analyzed in terms of cognitive domain levels. Indeed, emphasis is placed on a holistic approach and interdisciplinary approach in skill acquisition and permanent learning (MoNE, 2024b). Moreover, in the themes of the Turkish language curriculum, social studies emerges as the most prominent other discipline area in the themes at 5th, 6th, and 7th grade levels in the interdisciplinary relations sections under the heading of inter-program components (MoNE, 2024b, p. 87, 93, 98, 103, 108, 119, 125, 131, 137, 147, 159, 165, 171, 183). In this direction, when examining the social studies curriculum, it is seen that many writing activities are included such as introductory writings in line with learning evidence (MoNE, 2024a, p. 18), writing poetry and stories (MoNE, 2024a, p. 57-64-73-90), writing petitions (MoNE, 2024a, p. 78), writing compositions (MoNE, 2024a, p. 90); within the scope of learning-teaching practices, explanatory writings (MoNE, 2024a, p. 15), book cover design and introduction writing (MoNE, 2024a, p. 20), letter writing (MoNE, 2024a, p. 27), interpretive text writing (MoNE, 2024a, p. 104), proposition writing (MoNE, 2024a, p. 114); within the scope of enrichment applications, writing game rules (MoNE, 2024a, p. 24), story writing (MoNE, 2024a, p. 28-92-101), poetry writing (MoNE, 2024a, p. 50), column writing (MoNE, 2024a, p. 55), speech text writing (MoNE, 2024a, p. 55), scenario writing (MoNE, 2024a, p. 55-114),

digital comic writing (MoNE, 2024a, p. 88). This situation is an indication that special importance is given to writing skills in social studies lessons as well.

It is stated that in the social studies course, which has a rich content in terms of using writing activities for learning purposes, topics should be concretized, students should be actively involved in the learning process, and writing activities should be included for concept teaching (Ay and Başibüyük, 2022). It is expressed that the use of writing activities for learning purposes within the scope of social studies course will contribute to the development of students' writing skills, as well as their cognitive and communication skills (Karaçağıl, 2014). Kayaalp (2020) emphasizes the importance of writing skills in terms of knowledge, skill, and concept teaching, and states that the use of writing activities within the scope of social studies course will positively contribute to students' acquisition of targeted skills. In his study, Yasul (2019) determined that writing activities in social studies lessons positively affected academic achievement. In this context, it is considered important to examine the learning outcomes related to writing skills in the social studies curriculum (MoNE, 2024a), which is a closely related discipline to the Turkish course in terms of holistic development, from a cognitive perspective. At this stage, it was first determined which learning outcomes and process components related to writing skills are included in the curricula. Then, the identified learning outcomes and process components were interpreted according to the cognitive strategy area in which they are found. In interpreting the learning outcomes of the social studies course according to the cognitive domain levels in the context of writing skills, the explanations for the implementation of the relevant learning outcomes were examined, and it was determined whether these explanations were included in the field of writing skills. The questions sought to be answered during the research process can be listed as follows:

1. How are the writing skill-focused learning outcomes of the Middle School Turkish course curriculum distributed in terms of cognitive domain levels?
2. How are the writing skill-focused learning outcomes of the Social Studies course curriculum distributed in terms of cognitive domain levels?

METHOD

Qualitative research methods were used in this study. During the research process, the Middle School Turkish Language Curriculum (MoNE, 2024b) and Social Studies Curriculum (MoNE, 2024a) were examined in terms of learning outcomes related to writing skills. In determining the learning outcomes for writing skills, those directly related to writing skills in the Turkish language course were interpreted within the scope of cognitive domain levels. In the field of social studies, the explanations for learning outcomes were examined to determine whether they were related to writing skills. Following these assessments, the learning outcomes related to writing skills within the scope of the social studies course were examined in terms of cognitive domain levels. The stages of the research process are explained in Figure 1.

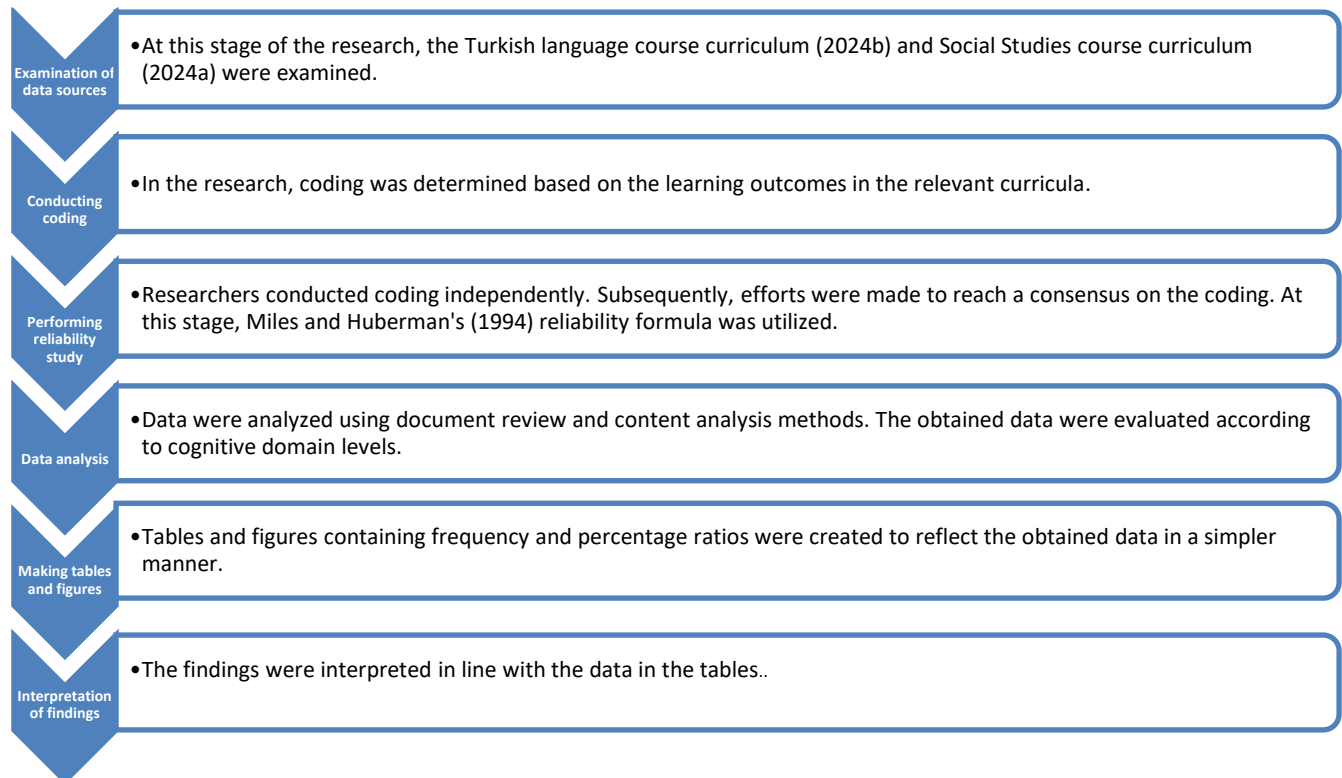


Figure 1. Research process

Data Collection

The Middle School Turkish Language Curriculum (MoNE, 2024b) and Social Studies Curriculum (MoNE, 2024a) were identified as data collection tools during the research process. In the examination of these curricula regarding writing skills, the learning outcomes and process components in this skill area were identified and analyzed. Document analysis was used in collecting data at this stage. Indeed, this method is used in the analysis of written materials related to the researched subject area (Yıldırım & Şimşek, 2008). At this stage, data for each curriculum were identified within its own scope. In identifying the writing skill learning outcomes within the Middle School Turkish Language Curriculum (MoNE, 2024b), writing domain skills for grades 5, 6, 7, and 8 were examined. In this context, the learning outcomes and process components of writing domain skills for each grade level were examined. Within the scope of the Social Studies Curriculum (MoNE, 2024a), learning outcomes and process components in the themes at each grade level were identified. Then, to determine whether the relevant learning outcomes and process components were related to writing skills, the learning-teaching practices for that learning outcome were examined. Accordingly, learning outcomes and process components identified as related to writing skills were determined as data sources. In this direction, a total of 32 learning outcomes within the social studies course were identified as being within the scope of writing skills.

Data Analysis

During the research process, content analysis was utilized in the data analysis phase. In this stage, learning outcomes related to writing skills were coded in the learning theme area in accordance with the relevant analysis method, and interpretations were made after theme-related arrangements (Yıldırım & Şimşek, 2008). Following this method, the aim was to determine which cognitive domain level each learning outcome belonged to. Indeed, cognitive learning is also included in the programs. In interpreting the learning outcomes according to cognitive strategies during the research process, the cognitive strategies in Bloom's revised taxonomy were used as a basis (Krathwohl, 2002). These cognitive processes are presented as the basic dimensions (strategies) of "remember, understand, apply, analyze, evaluate, create" (Krathwohl, 2002, p. 215). Accordingly, the learning outcomes included in the Middle School Turkish Language Curriculum (MoNE, 2024b) and Social Studies Curriculum (MoNE, 2024a) were categorized separately by the researchers. Subsequently, the researchers' determination of which cognitive strategy each learning outcome corresponded to was examined using Miles and Huberman's (1994) formula: $\text{Reliability} = \frac{\text{Agreement}}{\text{Agreement} + \text{Disagreement}} \times 100$. It was concluded that there was 90% agreement among researchers in their determinations of the cognitive domains in which the learning outcomes were placed. The results obtained were presented in tables and graphs for each course. Holistic interpretations were made by determining which cognitive strategy area each learning outcome belonged to. During interpretation, the sequence numbers of learning outcomes and process components in the curriculum were used as codes. The frequency of process components was expressed as frequency (f) and percentage (%). In the analyses within the Turkish language curriculum, interpretations were made regarding the learning outcomes. In the analyses of the social studies curriculum, the relevant data were interpreted in terms of frequency and percentage. The cognitive process steps in the tables and figures were based on Krathwohl (2002).

FINDINGS

In this section of the research, the learning outcomes for writing skills in the Turkish language course curriculum and the writing skill-focused learning outcomes in the social studies course curriculum were examined according to the cognitive domain levels of the revised Bloom's taxonomy. The Turkish language course curriculum and social studies course curriculum were analyzed according to cognitive domain levels, and the results obtained are presented in Table 1, 2 3, 4, and Figure 2 and Figure 3. The following abbreviation was used in the table for the Turkish course, showing learning outcomes and process components. For example, the abbreviation 1a refers to the process component (coded a) "Determines the writing topic, target audience, writing tool and environment, including digital, by analyzing purpose, context, and possibilities" of the learning outcome "Ability to manage the writing process," which is the first acquisition in the writing skill area for all grade levels (MoNE, 2024b: 56). In the table for social studies, the abbreviation (SB) was taken as it appears in the relevant program, and the numbers of the learning outcomes are given in the table. The findings regarding the analysis of writing skill-focused learning outcomes in the Turkish language course curriculum are presented in Table 1.

Table 1. Analysis of writing skill-focused learning outcomes in the Turkish language course curriculum according to cognitive processes

Cognitive Process Basic Dimension	Cognitive Process Sub-Dimension	Learning Outcomes (T.Y.5, 6, 7, 8)	Frequency (f)	Total
Remembering	Recognizing	16a	1	2
	Recalling	4a	1	
Understanding	Interpretation	-	-	13
	Exemplification	-	-	
	Classification	9a, 10a, 12b, 12c, 18b	5	
	Summarization	16b	1	
	Inference	8b, 18a	2	
	Comparison	9b, 9c, 13c, 17a, 18d	5	
	Explanation	-	-	
Application	Executive	2a, 2b, 3a, 3b, 6a, 6b	6	21
	Implementation	1b, 4c, 6c, 7c, 9d, 10d, 17b, 18c, 19a, 19b, 20a, 20b, 21a, 21b, 22c	15	
Analysis	Decomposition	1a, 5a, 10b, 11a, 12a, 16d	6	12
	Organization	1c, 5b, 10c, 11b	4	
	Association	4b, 8a	2	
Evaluation	Checking	13a, 14a, 14b, 15a, 22a, 22b	6	11
	Critiquing	5c, 8c, 13b, 14c, 15b	5	
Creating	Generating	7a, 7b, 7d, 11c, 12d, 13d, 15c, 16e	8	9
	Planning	-	-	
	Producing	16c	1	
Total				68

When examining the Middle School Turkish Language Curriculum (MoNE, 2024), there are a total of twenty-two learning outcomes and sixty-seven process components for the writing skills area. These learning outcomes and process components have been analyzed based on the cognitive levels in Bloom's revised taxonomy. The following findings were reached when examining the process components of learning outcomes according to the basic levels of the cognitive domain. Remember (f:2), Understand (f:13), Apply (f:21), Analyze (f:12), Evaluate (f:11), Create (f:9). When these findings are examined according to sub-dimensions, it has been determined that the process components of learning outcomes are mostly directed towards the application sub-dimension (f:15) of the apply basic level. It was concluded that eight process components are directed towards the produce sub-dimension (f:8) of the create basic level, six process components each are directed towards the check sub-dimension of the evaluate level, the execute sub-dimension of the apply level, and the differentiate sub-dimension of the analyze level (f:6). Five process components each are directed towards the critique sub-dimension of the evaluate basic level, and the classify and compare sub-dimensions of the understand basic level (f:5). Four learning outcomes are directed towards the organize sub-dimension of the analyze basic level (f:4), two process components are directed towards the infer sub-dimension of the understand basic level and the attribute sub-dimension of the analyze basic level (f:2). It has been determined that one learning outcome each is directed towards the recognize and recall sub-dimensions of the remember basic level, the summarize sub-dimension of the understand basic level, and the generate sub-dimension of the create basic level (f:1).

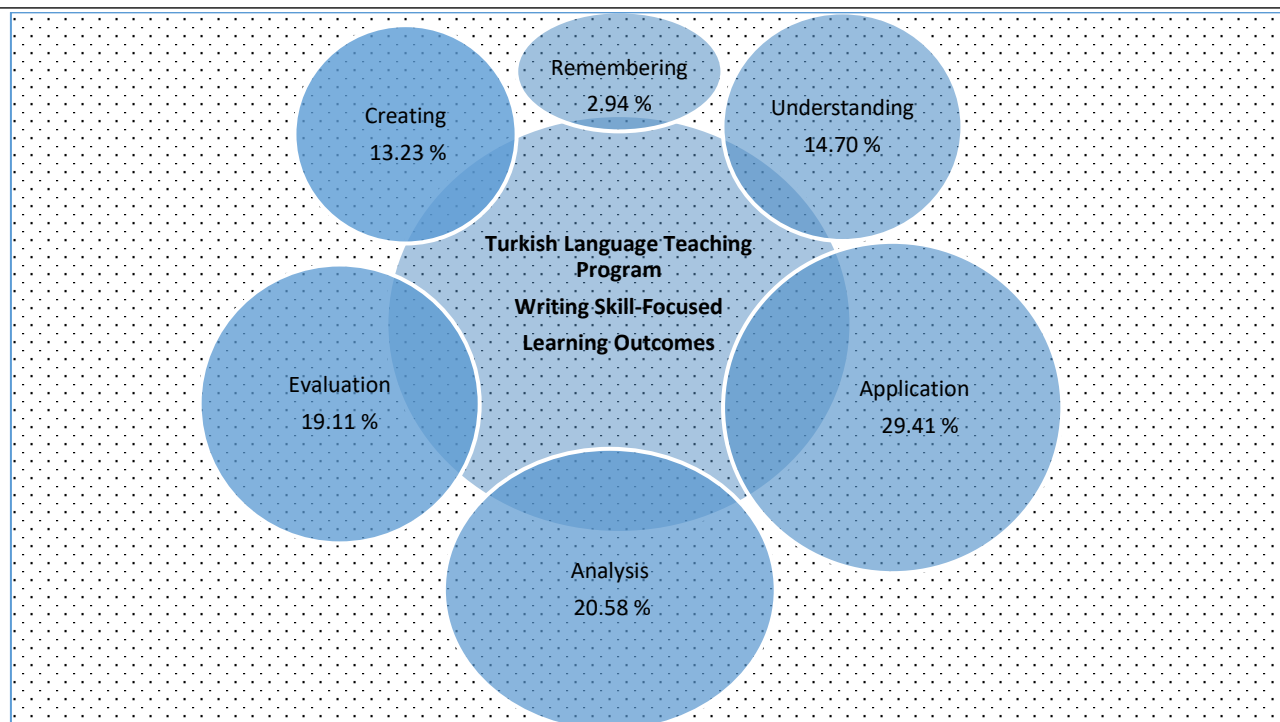


Figure 2. Distribution of writing skill-focused learning outcomes in the Turkish language teaching curriculum according to cognitive processes

When examining the learning outcomes for the writing skill area in the Turkish language teaching curriculum in terms of proportional distribution according to the basic cognitive domain steps shown in Figure 2, it is observed that the skill area with the most writing learning outcomes is application (29.41%). The analyze step stands out as the cognitive domain with the second highest number of learning outcomes (20.58%). Similarly, learning outcomes related to the evaluate step are also a more emphasized area in the curriculum (19.11%). Learning outcomes for the create step, which is the highest level cognitive domain, are relatively less preferred (13.23%). It is seen that learning outcomes for the understand step are the fifth most included learning area (14.7%). When writing learning outcomes are evaluated in terms of cognitive domain, the remember step is determined as the least preferred cognitive domain (2.94%).

Table 2. Distribution of learning outcomes in the writing skill area of the Turkish language curriculum by grade levels

Cognitive ProcessBasic Dimension	Grade			
	Grade 5 (f)	Grade 6 (f)	Grade 7 (f)	Grade 8 (f)
Remembering	2	2	2	2
Understanding	13	13	13	13
Application	21	21	21	21
Analysis	12	12	12	12
Evaluation	11	11	11	11
Creating	9	9	9	9
Total	68	68	68	68

When examining the writing skill-focused learning outcomes in the Turkish language curriculum in terms of grade levels through Table 2, it is observed that the learning outcomes and process components for 5th, 6th, 7th, and 8th-grade levels are the same for each grade level, although there are some changes in the subject area scope. In this respect, it can be stated that the learning outcomes and process components of the writing skill area show an equal distribution in terms of grade levels within the scope of the Turkish language curriculum. Findings related to the analysis of writing skill-focused learning outcomes in the social studies curriculum are presented in Table 3.

Table 3. Analysis of writing skill-focused learning outcomes in the social studies curriculum according to cognitive processes

Cognitive Process Basic Dimension	Cognitive Process Sub-Dimension	Learning Outcomes (SB. 5, 6, 7)	Frequency (f)	Total
Remembering	Recognizing	-	-	-
	Recalling	-	-	
Understanding	Interpretation	4.1.1; 4.4.1; 4.4.2; 5.2.2. 6.1.2; 6.4.2; 7.3.2; 7.5.1	8	19
	Exemplification	-	-	
	Classification	4.2.3; 6.5.2	2	
	Summarization	5.5.3; 6.3.4; 7.4.1	3	
	Inference	4.6.1; 5.4.2; 6.1.1; 7.6.1; 7.6.2	5	
	Comparison	4.3.1	1	
	Explanation	-	-	
Application	Executive	5.4.4	1	1
	Implementation	-	-	
Analysis	Decomposition	5.4.1	1	2
	Organization	-	-	
	Association	5.1.1	1	
Evaluation	Checking	7.1.1; 7.3.1	2	5
	Critiquing	5.3.2; 5.4.3; 5.6.1	3	
Creating	Generating	4.1.3; 4.4.3; 7.1.3	3	5
	Planning	6.3.2; 7.3.3	2	
	Producing	-	-	
Total				32

When examining the Social Studies Curriculum (MoNE, 2024a) in terms of learning outcomes related to writing skills as shown in Table 3, it is observed that there are a total of thirty-two learning outcomes. These learning outcomes have been analyzed based on the cognitive levels in Bloom's revised taxonomy. During the analysis phase, attention was paid to the application dimension of each learning outcome. For example, regarding the learning outcome "SB.5.5.3. Ability to summarize economic activities in the province where one lives," when examining the explanations in the teaching-learning practices, students are expected to present their thoughts in writing (MoNE, 2024a: 59). Accordingly, it was observed that this learning outcome is related to the summarizing sub-dimension of the understand level, and the research data was created using this method. When examining the learning outcomes related to writing skills in social studies according to the basic levels of the cognitive domain, the following results were obtained: Understand (f:19), Apply (f:1), Analyze (f:2), Evaluate (f:5), Create (f:5). When these results are examined according to sub-dimensions, it is seen that the learning outcomes are most frequently related to the interpreting sub-dimension of the understand level (f:8). Five learning outcomes were found to be related to the inferring sub-dimension of the understand level (f:5). It was determined that three learning outcomes each were related to the summarizing sub-dimension of the understand level, the critiquing sub-dimension of the evaluate level, and the generating sub-dimension of the create level (f:3). Two learning outcomes each were found to be related to the classifying sub-dimension of the understand level, the checking sub-dimension of the evaluate level, and the planning sub-dimension of the create level (f:2). One learning outcome each was found to be related to the comparing sub-dimension of the understand level, the executing sub-dimension of the apply level, and the differentiating and organizing sub-dimensions of the analyze level (f:1).

Table 4. Distribution of writing skill-focused learning outcomes in the social studies curriculum by grade levels

Cognitive Process Basic Dimension	Grade			
	Grade 4 (f)	Grade 5 (f)	Grade 6 (f)	Grade 7 (f)
Remembering	-	-	-	-
Understanding	6	3	5	5

Cognitive Process	Grade			
	Grade 4 (f)	Grade 5 (f)	Grade 6 (f)	Grade 7 (f)
Application	-	1	-	-
Analysis	-	2	-	-
Evaluation	-	3	-	2
Creating	2	-	1	2
Total	8	9	6	9

When examining the writing skill-focused learning outcomes in the social studies curriculum in terms of grade levels through Table 4, it can be said that there is a similarity in the distribution of writing skill-focused learning outcomes across grade levels, but there are relatively fewer learning outcomes for writing skills at the sixth-grade level. At the fourth-grade level, it is observed that (f=6) learning outcomes are directed towards the basic step of understanding, while (f=2) learning outcomes are directed towards the basic step of creating. At the fifth-grade level, it has been determined that (f=3) learning outcomes are directed towards understanding, (f=1) learning outcome towards applying, (f=2) learning outcomes towards analyzing, and (f=3) learning outcomes towards evaluating. At the sixth-grade level, (f=5) learning outcomes are directed towards understanding, while (f=1) learning outcome is directed towards creating. At the seventh-grade level, (f=5) learning outcomes are directed towards understanding, (f=2) learning outcomes towards evaluating, and (f=2) learning outcomes towards creating. This situation shows that the learning outcomes for writing skills in the social studies course are mostly concentrated in the basic step of understanding. In this respect, higher-level skill steps related to the research topic are considered as an area that can be developed.

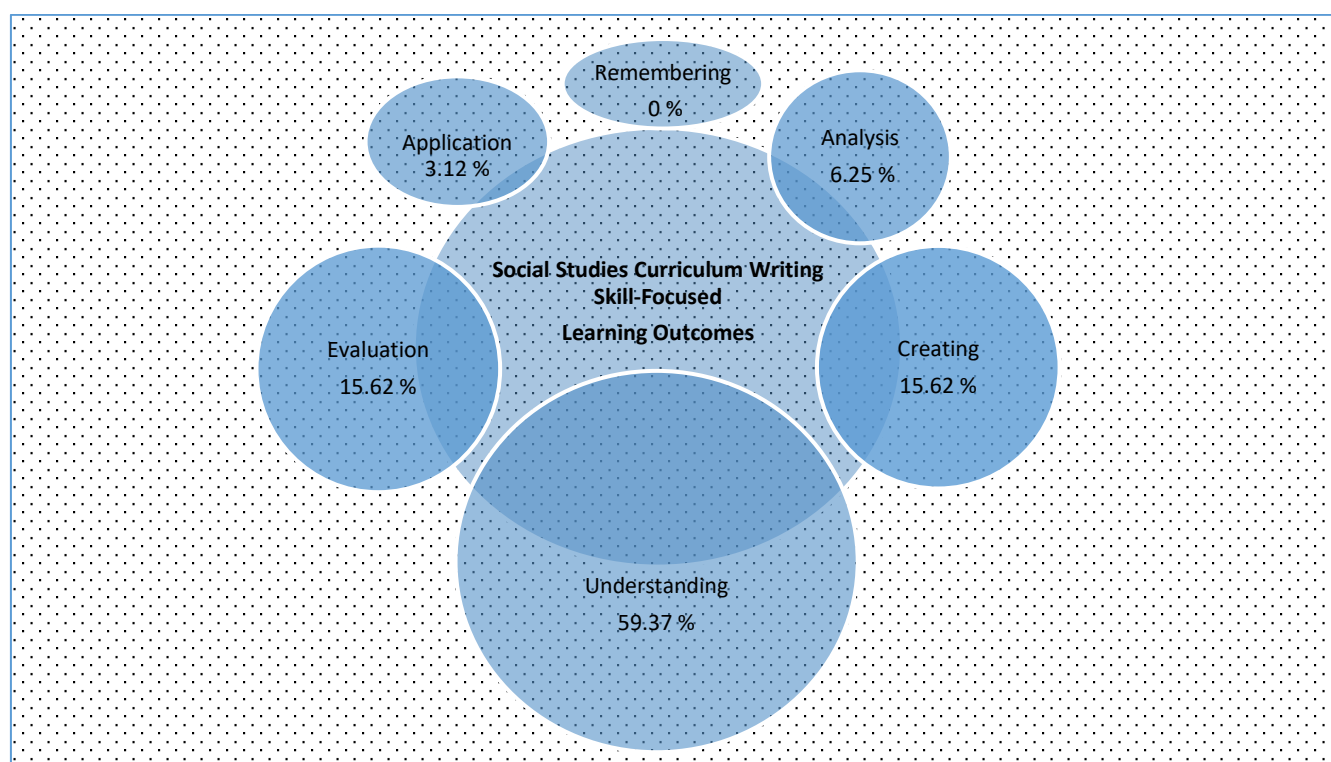


Figure 3. Distribution of writing skill-focused learning outcomes in the social studies curriculum according to cognitive processes

When examining the learning outcomes for the writing skills area in the social studies curriculum shown in Figure 3 in terms of proportional distribution according to the basic cognitive domain levels, it is observed that the skill area with the most writing learning outcomes is the basic level of understanding (59.37%). The basic levels of evaluate and create were found to have the second highest number of learning outcomes for the writing skills area (15.62%). In the relevant curriculum, it was determined that the basic level of analyze (6.25%) and the basic level of apply (3.12%) encompassed fewer learning outcomes compared to other cognitive domains. No learning outcome related to writing skills was identified for the basic level of remember.

DISCUSSION, CONCLUSION AND RECOMMENDATIONS

Curricula stand out as one of the primary reference sources in conducting lessons, from the preparation of textbooks to the creation of lesson plans. This situation ensures that curricula play a significant role in developing writing skills in addition to other skills. It is emphasized that writing in education should be viewed as a process rather than a tool (McVey, 2008). Curricula have a significant impact on developing writing skills, which is a process. Indeed, it is emphasized that improving academic writing instruction is important in the process of developing curricula (Paltridge, 2004). In this context, various studies on writing education in Turkish language teaching curricula have been observed (Güzel & Karadağ, 2013; Temizkan & Atasoy, 2014). Güzel and Karadağ (2013) point out that speaking and writing outcomes in the Turkish language curriculum are the same across grade levels, which they consider problematic. When evaluating the results, it is seen that the process components of learning outcomes in the Turkish language curriculum are mostly oriented towards the application level and higher-order cognitive domain levels such as analyze and evaluate, although there is a close distribution in terms of lower and higher-order cognitive domain levels. The presence of an application dimension in the writing education process is important for the development of students' writing skills. Supporting these skills with higher-order cognitive strategies can positively affect student achievement. In this respect, it is recommended to focus more on the sub-dimensions of creating, planning, and producing, especially in the create basic level of outcomes. In the learning outcomes related to the writing domain in the social studies curriculum, it is observed that the understand level stands out significantly compared to other skill areas. When the learning outcomes for the understand and apply levels are evaluated together, it can be stated that the learning outcomes for the writing skill area are more oriented towards the lower-order cognitive domain in proportional terms. However, it is also seen that learning outcomes related to writing skills oriented towards higher-order cognitive domains hold an important position in the curriculum. Indeed, it can be stated that the cognitive domain levels of analyze, evaluate, and create account for nearly 40% of the learning outcomes related to writing skills. A balanced distribution of cognitive skill areas in curricula is considered important. In their research on this subject, Sarıoğlu and Babadoğan (2023) state that teachers believe writing outcomes in the Turkish language curriculum should be oriented towards all dimensions of mental processes.

In addition to the inclusion of cognitive processes in the program, it is also important that these strategies can be used by students. This is because the appropriate use of strategies for learning purposes is considered significant in academic success. Belet and Yaşar (2007) found in their research on this subject that the use of learning strategies in Turkish lessons contributed to the development of students' reading and writing skills. There are also studies on teaching programs in terms of writing skills. Tabak and Göçer (2013) state that the Turkish course curriculum they examined carries the characteristics of a process-oriented writing approach. Temizkan and Atasoy (2014) examined the Turkish teaching programs of the Republican period in the context of writing skills. Attention is drawn to the process-based approach in developing writing skills, and particular emphasis is placed on the importance of the preparation dimension of this approach (Doğan, 2020). Furthermore, feedback is considered important in the writing process. It is stated that providing feedback in writing processes will contribute to the improvement of text quality (Wischgoll, 2017). It is noted that teachers providing feedback to students during the writing process can contribute to the development of writing skills (Wulandari, 2022). Although writing is a language skill, it stands out as an important competence not only in the field of language education but also in other disciplines, as it is a means for individuals to express themselves. In this direction, the development of students' language skills, including writing, is considered a process with multifaceted effects. Indeed, as a result of the study conducted by Kayaalp (2020), it was concluded that learning-oriented writing activities in social studies lessons have positive effects on students' cognitive and affective learning.

Within the scope of the Turkish language course curriculum, it has been determined that the learning outcomes and process components of the writing skills area are equally distributed across grade levels. When examining the writing skill-focused learning outcomes in the social studies course curriculum in terms of grade levels, it has been concluded that the learning outcomes for writing skills in the social studies course are mostly concentrated at the basic level of comprehension. In this respect, higher-level skill stages related to the subject of the research are considered as an area that can be developed.

In the context of the results obtained, it is considered important to have a balanced distribution of learning outcomes in the field of writing skills, which is one of the self-expression skills of students, in terms of cognitive domain steps in the relevant curriculum. It is thought that reflecting this situation in the curriculum will positively affect students' written expression skills. This research is limited to examining the writing skill-focused learning outcomes included in the curriculum. This is considered important in terms of allowing a detailed examination of writing skill-focused learning outcomes according to cognitive process dimensions. Based on the results obtained in this research, recommendations have been made.

- It is recommended to include more learning outcomes related to higher-order cognitive domain skills and especially the basic step of analysis in the written expression of students' research skills and the data they obtain.
- It is suggested to examine the curricula from an interdisciplinary approach in terms of other language skills such as listening, reading, and speaking.
- It is recommended to conduct practical studies on how the learning outcomes in the programs related to writing skills are reflected in students' texts.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, author-ship, and/or publication of this article.

Statements of publication ethics

We hereby declare that the study has not unethical issues and that research and publication ethics have been observed carefully.

Author contribution statements

All authors wrote the introduction, methods sections, discussed the results, and contributed to the final manuscript.

Researchers' contribution rate

The study was conducted and reported with equal collaboration of the researchers.

Ethics Committee Approval Information

This study is based on the examination of Turkish language and Social Studies curricula available on the Ministry of National Education's website. Therefore, it is not subject to Ethics Committee Approval.

REFERENCES

- Ay, A., & Başbüyük, A. (2022). Sosyal bilgiler dersinde öğrenme amaçlı yazma etkinlikleri kullanımına ilişkin öğrenci deneyimleri. *İnönü Üniversitesi Eğitim Fakültesi Dergisi*, 23(2), 1112-1131. <https://doi.org/10.17679/inuefd.1146311>
- Belet, Ş. D. & Yaşar, Ş. (2007). Öğrenme stratejilerinin okuduğunu anlama ve yazma becerileri ile Türkçe dersine ilişkin tutumlara etkisi. *Eğitimde Kuram ve Uygulama*, 3(1), 69-86.
- Cer, E. (2019). The instruction of writing strategies: The effect of the metacognitive strategy on the writing skills of pupils in secondary education. *SAGE Open*, 9(2). <https://doi.org/10.1177/2158244019842681>
- Coşkun, H. & Narinç, F. N. (2018). 2017 Türkçe öğretim programı esas alınarak hazırlanan 5. sınıf Türkçe ders kitabında yer alan yazma ve konuşma etkinliklerinin incelenmesi. *Iğdır Üniversitesi Sosyal Bilimler Dergisi*, 16, 409-427.
- Doğan, Y. (2020). Süreç temelli yazma eğitiminin bir boyutu olarak "hazırlık". *Ana Dili Eğitimi Dergisi*, 8(2), 326- 342.
- Gündüz, O. & Şimşek, T. (2016). *Uygulamalı yazma eğitimi el kitabı* (3. Baskı). Grafiker Yayınları.
- Güzel, A. & Karadağ, Ö. (2013). Anlatma becerileri açısından "Türkçe dersi öğretim programı (6, 7, 8. sınıflar)"na eleştirel bir bakış. *Ana Dili Eğitimi Dergisi*, 1(1), 45-52.
- Karacaoğlu, M.Ö., Dağ, M. & Uzun, O. (2021). Türkçe ders kitaplarındaki yazma etkinliklerinin incelenmesi. *IBAD Sosyal Bilimler Dergisi*, (10), 96-126.
- Karaçağıl, C. (2014). *Sosyal bilgiler derslerinde öğrenme amaçlı yazma etkinlikleri kullanımının öğrenci akademik başarısına etkisi* [Yayımlanmamış yüksek lisans tezi]. Muğla Sıtkı Koçman Üniversitesi.
- Karagöl, E., Tarakcı, R. (2018). Yazma becerisi kazanımları açısından ortaokul türkçe ders kitapları. *International Journal of Languages' Education and Teaching*, 6(4), 87-100.
- Kayaalp, F. (2020). *Sosyal bilgiler dersinde öğrenme amaçlı yazma etkinliklerini kullanmanın bilişsel ve duyuşsal öğrenmelere etkisi* [Yayımlanmamış doktora tezi]. Atatürk Üniversitesi.
- Kellogg, R. T. (2008). Training writing skills: A cognitive developmental perspective. *Journal of Writing Research*, 1(1), 1-26.
- Kellogg, R. T., & Raulerson, B. A. (2007). Improving the writing skills of college students. *Psychonomic Bulletin & Review*, 14 (2), 237-242.
- Klimova, B. F. (2014). Approaches to the teaching of writing skills. *Procedia-Social and Behavioral Sciences*, 112, 147-151.
- Krathwohl, D. R. (2002). A revision of Bloom's taxonomy: An overview. *Theory Into Practice*, 41(4), 212-218, DOI: 10.1207/s15430421tip4104_2
- McVey, D. (2008). Why all writing is creative writing. *Innovations in Education and Teaching International*, 45(3), 289-294. <https://doi.org/10.1080/14703290802176204>
- Ministry of National Education [MoNE]. (2006). *İlköğretim Türkçe dersi (6, 7, 8. Sınıflar) öğretim programı*. Millî Eğitim Bakanlığı.
- Ministry of National Education [MoNE]. (2019). *Türkçe dersi öğretim programı (İlkokul ve ortaokul 1, 2, 3, 4, 5, 6, 7 ve 8. sınıflar)*. Millî Eğitim Bakanlığı.
- Ministry of National Education [MoNE]. (2024a). Sosyal bilgiler dersi öğretim programı (4, 5, 6 ve 7. sınıflar). Erişim Adresi: <https://tymm.MoNE.gov.tr/upload/program/2024programsos4567Onayli.pdf>
- Ministry of National Education [MoNE]. (2024b). *Ortaokul Türkçe dersi öğretim programı (5, 6, 7 ve 8. sınıflar)*. Erişim Adresi: <https://tymm.MoNE.gov.tr/upload/program/2024programtur5678Onayli.pdf>
- Miles, M. B., & Huberman, A. M. (1994). *Qualitative data analysis* (2nd Ed.). Thousand Oaks, Sage.
- Paltridge, B. (2004). Academic writing. *Language Teaching*, 37(2), 87-105. doi:10.1017/S0261444804002216

- Rao, Z. (2007). Training in brainstorming and developing writing skills. *ELT Journal*, 61(2), 100-106.
- Sarioğlu, S. & Babadoğan, M. C. (2023). *Türkçe Dersi öğretim programı'nın yazma beceri alanının değerlendirilmesi* [Bildiri Sunumu]. International Education Congress, Ankara.
- Tabak, G. & Göçer, A. (2013). 6-8. sınıflar Türkçe dersi öğretim programının ürün ve süreç odaklı yazma yaklaşımları çerçevesinde değerlendirilmesi. *Ahi Evran Üniversitesi Kırşehir Eğitim Fakültesi Dergisi (KEFAD)*, 14 (2), 147-169.
- Temizkan, M. & Atasoy, A. (2014). Cumhuriyet dönemi Türkçe eğitimi ve öğretimi programlarında yazma becerisi. *Ana Dili Eğitimi Dergisi*, 2(2), 42-71.
- Ungan, S. (2007). Yazma becerisinin geliştirilmesi ve önemi. *Erciyes Üniversitesi Sosyal Bilimler Enstitüsü Dergisi*, 23 (2), 461-472.
- Wischgoll, A. (2017). Improving undergraduates' and postgraduates' academic writing skills with strategy training and feedback. *Frontiers in Education* (2). 1-15. <https://doi.org/10.3389/feduc.2017.00033>
- Wulandari, Y. (2022). Effective feedback to improve students' writing skills. *English Education, Linguistics, and Literature Journal*, 1(1), 10-17.
- Yasul, A. F. (2019). *Sosyal bilgiler dersinde kullanılan öğrenme amaçlı yazma etkinliklerinin akademik başarıya, öğrenmenin kalıcılığına ve yazmaya yönelik tutuma etkisi* [Yayınlanmamış yüksek lisans tezi]. Atatürk Üniversitesi.
- Yıldırım, A. & Şimşek, H. (2008). *Sosyal bilimlerde nitel araştırma yöntemleri* (6. Baskı). Seçkin Yayıncılık.



| Research Article / Araştırma Makalesi |

Teacher Perceptions on the Century of Turkey Educational Model Based on The Functions of Education

Eğitimin İşlevleri Temelinde Türkiye Yüzyılı Maarif Modeline İlişkin Öğretmen Görüşleri

Esen BAĞÇIVAN¹, Hilal KAHRAMAN², Soner DOĞAN³

Keywords

1. Century of Türkiye Education Model
2. Functions of Education
3. Teacher Perspectives
4. Qualitative Research
5. Education Policy

Anahtar Kelimeler

1. Türkiye Yüzyılı Maarif Modeli
2. Eğitimin İşlevleri
3. Öğretmen Görüşleri
4. Nitel Araştırma
5. Eğitim Politikası

Received/Başvuru Tarihi

30.02.2025

Accepted / Kabul Tarihi

10.10.2025

Abstract

This study aims to evaluate the Century of Türkiye Education Model (CTEM) from the perspectives of teachers in terms of the individual, social, economic, and political functions of education. The research employed a qualitative design based on the phenomenological approach. Data were collected through semi-structured interviews with 50 teachers from various subject areas working in public schools in Sivas where CTEM is being implemented. Additionally, the official CTEM Core Text published by the Ministry of National Education was analyzed using document analysis. The data were examined through MAXQDA software and subjected to descriptive and content analysis. Findings reveal that CTEM strongly reflects the individual and social functions of education, while certain deficiencies remain in the practical implementation of its economic and political functions. Teachers reported that CTEM contributes positively to individual development, value transmission, and social integration, but emphasized the need for more concrete practices in areas such as economic competencies, career guidance, and citizenship education. The study aims to present the field-level implications of educational policy from the viewpoint of teachers and to identify both the strengths and areas in need of improvement within CTEM.

Öz

Bu araştırmanın amacı, Türkiye Yüzyılı Maarif Modeli'nin eğitimin bireysel, toplumsal, ekonomik ve siyasal işlevleri açısından öğretmen görüşlerine dayalı olarak değerlendirilmesidir. Araştırma, nitel desende ve olgubilim yaklaşımı doğrultusunda gerçekleştirilmiştir. Veriler, Sivas il merkezindeki devlet okullarında Maarif Modeli'ni uygulayan farklı branşlardan 50 öğretmenle yapılan yarı yapılandırılmış görüşmeler aracılığıyla toplanmıştır. Ayrıca, Millî Eğitim Bakanlığı tarafından yayımlanan Türkiye Yüzyılı Maarif Modeli Ortak Metni doküman analizi yöntemiyle incelenmiştir. Elde edilen veriler MAXQDA yazılımı kullanılarak betimsel ve içerik analizine tabi tutulmuştur. Araştırma bulguları, modelin eğitimin bireysel ve toplumsal işlevlerini güçlü biçimde yansıttığını; buna karşın ekonomik ve siyasal işlevlerin uygulamaya dönük bazı eksiklikler barındırdığını ortaya koymaktadır. Öğretmenler, modelin bireysel gelişim, değer aktarımı ve toplumsal uyum gibi alanlarda katkı sağladığını belirtirken; ekonomik yeterlilikler, mesleki yönlendirme ve vatandaşlık eğitimi gibi konularda daha somut uygulamalara ihtiyaç olduğunu vurgulamışlardır. Bu çalışma, eğitim politikalarının saha düzeyindeki yansımalarını öğretmen bakış açısıyla ortaya koymayı ve Maarif Modeli'nin güçlü yönleri ile geliştirilmesi gereken alanlarını belirlemeyi amaçlamaktadır.

¹ Corresponded Author, Sivas Cumhuriyet University, Faculty of Education, Department of Educational Sciences, Sivas, Türkiye; <https://0009-0003-4724-6972>

² Sivas Cumhuriyet University, Faculty of Education, Department of Educational Sciences, Sivas, Türkiye; <https://0000-0002-5680-8362>

³ Sivas Cumhuriyet University, Faculty of Education, Department of Educational Sciences, Sivas, Türkiye; <https://0000-0003-2013-3348>

INTRODUCTION

Education is a multidimensional process that not only enables individuals to acquire knowledge but also facilitates their integration into society, fosters their development into productive members, promotes the internalization of cultural values, and ensures the transmission of these values to future generations (Türer, 2006). Accordingly, the meaning attributed to education varies across historical, cultural, and ideological contexts. A review of educational definitions reveals that different scholars approach the concept from various dimensions. For instance, Tezcan (1981) defines education as a set of processes through which individuals develop their abilities, attitudes, and constructive behaviors within the society in which they live. Adem (1977), on the other hand, describes it as all forms of training activities aimed at cultivating intellectual and physical capacities in both children and adults. Koçer (1974) views education as the totality of processes through which individuals acquire positive attitudes and behaviors, while Tunç (1969) characterizes it as a social process in which individuals internalize societal standards, beliefs, and lifestyles. Dewey's perspective on education is grounded in the idea that individuals deepen their social structures by continually renewing their experiences and exercising control over their own learning processes (as cited in Varış, 1991). However, more contemporary definitions emphasize that education should not only be understood as knowledge transmission but also as a means of fostering 21st-century skills such as creativity, critical thinking, digital literacy, global awareness, and social responsibility (OECD, 2019). Collectively, these definitions underscore that education is not solely a pedagogical phenomenon but also a social, cultural, and political institution.

Durkheim, who conceptualizes education as a social institution, defines it as the process through which adult generations shape younger generations both intellectually and morally in preparation for social life (as cited in Ergün, 1987). In this respect, education serves not only as a vehicle for change but also as a mechanism for maintaining social equilibrium, since societies are in a constant state of transformation (Sönmez, 1991). The individual, social, economic, and political functions of education are designed to respond to this transformation (Doğan, 2015). It is through education that individuals develop their abilities, adapt to social norms, participate in the economic process as productive citizens, and embrace democratic values to engage in political life (Şişman, 2007; Hoşgörür & Taştan, 2012a). These functions of education—individual, social, economic, and political—are operationalized through educational programs and curricula in line with intended outcomes and societal needs, thereby translating the broader purposes of education into practice within daily educational experiences. Globally, countries have restructured their education systems in line with these changing functions and expectations. Finland prioritizes student well-being and learner autonomy, Singapore highlights technology-driven innovation, while Canada integrates multicultural values and equity-based learning approaches (Sahlberg, 2011; OECD, 2020). These examples demonstrate that education is increasingly seen as a transformative tool for preparing individuals to thrive in complex, rapidly evolving societies.

From a theoretical standpoint, education has been widely discussed through different perspectives. Human Capital Theory views education as an investment that enhances economic productivity and competitiveness (Becker, 1993), while Functionalist Theory emphasizes its role in social integration, cultural transmission, and role allocation (Durkheim, 2012). More recent perspectives such as Critical Pedagogy underline the emancipatory power of education in challenging inequalities and fostering democratic participation (Freire, 2000). In addition, critical economists argue that education also reproduces social and economic inequalities, as highlighted in the seminal work *Schooling in Capitalist America* (Bowles & Gintis, 1976). Incorporating these theoretical frameworks provides a foundation for analyzing Turkey's new educational reforms and situating them within broader global debates on the purposes and functions of education.

Achieving the desired profile of individuals and society in education is directly linked to the quality of curricula. Therefore, the transformation aimed for in education necessitates not only pedagogical innovations but also substantial changes in content. In this context, curricula function as fundamental tools that facilitate not only the delivery of course content but also the transmission of values, social integration, and economic participation. Thus, every educational reform movement is aligned not only with goals related to individual development but also with broader social and political objectives.

In this context, curriculum changes signify not merely a renewal of content but a comprehensive restructuring of the entire education system. In Turkey, earlier reforms include the constructivist-based curriculum reform of 2005 and the skills-based content updates introduced in 2017. Globally, countries such as Finland, Singapore, and Canada stand out with education systems that emphasize skills-based learning and account for individual differences (Sahlberg, 2011). In 2024, Turkey introduced a new paradigm shift in its education system with the launch of the Century of Türkiye Education Model (CTEM). CTEM is a comprehensive transformation project that aims to integrate an educational vision grounded in national and moral values with the demands of the modern era. The model encompasses curriculum development, teacher training, instructional material design, and assessment processes (Akpınar & Köksalan, 2024).

The aim of this study is to evaluate the extent to which the holistic development approach proposed by CTEM corresponds to the individual, economic, social, and political functions of education. The functions of education systems present a multifaceted structure, ranging from the integration of the individual into society and economic productivity to the development of political

consciousness and the transmission of cultural values (Şişman, 2007). CTEM is a comprehensive model that claims to address all of these functions. In this study, the degree to which the model reflects these functions has been assessed through both theoretical document analysis and the perceptions and experiences of practicing teachers regarding the model. This is especially important as the success of educational policies in practice is largely linked to the implementing role of teachers.

Studies in the literature concerning CTEM generally focus on the examination of curricula (Köçer, 2025; Hevedanlı & İnce, 2025; Banaz, 2024; Kuzu, Göçer & Akçay, 2024; Yurdakal, 2024; Üzümcü & Abanoz, 2024; Berk & Özer, 2024), the evaluation of the model itself (Ülçay, 2024; Karataş, 2024; Yıldırım & Çalışkan, 2024; Akpınar & Köksalan, 2024), analyses based on the different values and skills encompassed by the model (Doğan & Aşlamacı, 2025; Sezer & Dedeoğlu, 2024; Efe, 2024), the philosophical and ideological examination of the model (Arslankara & Arslankara, 2024; Turan & Nazıroğlu, 2024; Akpınar & Köksalan, 2024), and teacher perspectives regarding the model (Duyul, Duyul, Kesman & Kesman, 2025; Uygun & Akgül, 2024). This study distinguishes itself from the existing literature by investigating how and to what extent the holistic development approach proposed by CTEM reflects the individual, economic, social, and political functions of education, both theoretically and through the perceptions of practicing teachers. In this respect, the study aims to present a multidimensional evaluation of the model, supported not only at the curriculum level but also through teachers' classroom practices and experiences. Thus, it seeks to offer an original perspective on the field-level reflections of the model and contribute new insights to the literature.

METHOD/MATERIALS

Research Design

This study is a qualitative investigation aimed at evaluating CTEM based on the functions of education through both theoretical document analysis and the assessment of perceptions and experiences of practicing teachers regarding the model. The research design adopted a phenomenological approach. Phenomenology seeks to gain an in-depth understanding of individuals' experiences related to a particular phenomenon and the meanings they ascribe to those experiences (Yıldırım & Şimşek, 2022; Creswell, 2016). Accordingly, the study aimed to analyze the core text of the model and to evaluate teacher perspectives within the context of the functions of education.

Study Group

Data for this study were collected through document analysis and face-to-face interviews. Accordingly, the document analysis focused on CTEM Core Text of Curriculum Programs (2024), considered a fundamental document that systematically defines the model's holistic development approach and its connections to the individual, social, economic, and political functions of education. This provided a direct and detailed analysis of the model's theoretical foundations, thereby establishing a solid institutional framework for the practice-based examination conducted through teacher perspectives.

In the phase of the study where the extent to which CTEM reflects the individual, economic, social, and political functions of education was evaluated through the perceptions of practicing teachers, participants were selected using purposive sampling methods, specifically criterion sampling and maximum variation sampling. Criterion sampling is a strategy that involves including individuals who possess specific characteristics relevant to the study (Yıldırım & Şimşek, 2022). Maximum variation sampling aims to capture a wide range of experiences and situations representing different dimensions of the phenomenon under investigation (Patton, 2014). In line with these sampling strategies, specific criteria were established for selecting teachers to ensure diversity by including participants from different branches and grade levels. To this end, the criterion that teachers must be working in grades 1, 5, and 9, where CTEM has been implemented, was applied, and maximum variation was sought by involving teachers from various branches covered by the curriculum. Accordingly, the study group consisted of 50 teachers (see Table 1) working in public schools applying CTEM in the city center of Sivas, representing the branches of Turkish, Mathematics, Science, Social Studies, and Religious Culture and Ethics. The selection of these branches was made with attention to capturing the different dimensions of CTEM's holistic development approach by including subjects thought to represent both academic knowledge and value transmission within the curriculum.

Table 1. Demographic Characteristics of Participant

Branch	Gender	Years of Experience	Grade Level	Participant Codes
Mathematics	Female: 3, Male: 2	5–20 years	Middle School	T1, T2, T3, T4, T5
Turkish	Female: 6, Male: 8	5–21 years	Middle School	T6–T19
Science	Female: 1	16 years	Middle School	T20
Social Studies	Female: 1, Male: 3	12–19 years	Middle School	T21–T24
Classroom Teacher	Female: 3, Male: 8	1–33 years	Primary School	T25–T35
Religious Culture and Ethics	Male: 2	27 years	High School	T36, T37
Vocational Course	Male: 3	32–38 years	High School	T38–T40
Visual Arts	Male: 1	12 years	High School	T41

Branch	Gender	Years of Experience	Grade Level	Participant Codes
Information Technology	Female: 1, Male: 1	8–20 years	Middle School	T42, T43
Turkish Language and Literature	Female: 1, Male: 1	15 years	High School	T44, T45
Music	Male: 1	6 years	Middle School	T46
Philosophy	Male: 1	26 years	High School	T47
Geography	Male: 1	3 years	High School	T48
History	Male: 1	15 years	High School	T49
Biology	Female: 1	14 years	High School	T50

A total of 50 teachers participated in the study. Among them, 14 were Turkish language teachers, 11 were classroom teachers, 5 taught mathematics, 4 taught social studies, 3 were vocational course instructors, and 2 taught religious culture. Teachers from other branches participated individually. The gender distribution was generally balanced, with both female and male teachers represented across most branches. Participants' years of experience ranged from 1 to 38 years, with a significant portion having 10 or more years of experience. Teachers were employed at grades 1, 5, and 9, thereby ensuring diversity across the grade levels where the model is implemented.

Data Collection

In the data collection process, two qualitative research methods—document analysis and interview technique—were used together. In the first phase, the *Century of Türkiye Education Model* core text was examined in detail, and statements regarding the individual, social, economic, and political functions of education were coded and analyzed. In the second phase, semi-structured face-to-face interviews were conducted with teachers working in grades 1, 5, and 9, where the model was implemented. During the interviews, teachers were asked questions about the possible effects of the model on the education system and the extent to which the model fulfilled the functions of education and from which aspects. The interviews were conducted outside of school hours at times determined by the participants so as not to disrupt teaching and learning, and with participants' consent, audio recordings were taken. Each interview lasted approximately 20-25 minutes, and participants were identified by codes such as T1, T2 to ensure confidentiality. Afterwards, the recordings were transcribed into Word documents, and the data obtained from the interviews were comparatively analyzed with the findings of the core text analysis and systematically presented in the form of graphs and tables.

Validity and Reliability

To ensure the trustworthiness of the study, the criteria of credibility, dependability, transferability, and confirmability were taken into account (Lincoln & Guba, 1985). Various strategies were employed to ensure credibility, transferability, dependability, and confirmability. To ensure credibility, the strategy of triangulation was applied: data sources were diversified by analyzing both the core text of the *Century of Türkiye Education Model* and conducting semi-structured interviews with teachers from different branches. For transferability, a detailed description strategy was employed, presenting each stage (purpose, method, participant characteristics, findings) to the reader in detail. Furthermore, direct quotations obtained from the interviews and emphases in the core text were presented comparatively, thereby strengthening transferability. To ensure dependability, the strategy of researcher triangulation was used during data analysis. In this context, all three researchers actively participated in coding and interpreting the data; the codes, categories, and themes that emerged were constantly compared through the perspectives of different researchers. In addition, the use of MAXQDA software in the coding process provided an objective framework for data analysis. In this way, subjective judgments were reduced, and the data were interpreted more objectively and consistently. Finally, for confirmability, research auditing was employed. The conceptual framework, purpose, method, data collection tools, data analysis, and reporting stages of the study were evaluated in line with expert opinions, and deficiencies that emerged during this process were identified and necessary revisions were made.

Data Analysis

The analysis of the data obtained in the study was conducted using the content analysis method, which is appropriate for qualitative research designs. Content analysis is defined as “a research method that enables the systematic, objective, and quantitative or qualitative analysis of verbal, written, or visual messages conveyed through communication tools” (Krippendorff, 2004). In the first phase of analysis, the core text of the *Century of Türkiye Education Model* was examined. It was read from a holistic perspective in line with the research questions, and meaningful statements, sentences, or paragraphs were coded within the framework of the functions of education. To increase the reliability of the document analysis, the obtained data were continuously examined through a comparative approach, and coding consistency was ensured. In addition, new codes and themes that emerged during the analysis were interpreted within the holistic framework of the research objectives and the four basic functions of education (individual, social, economic, political). Later, codes carrying similar meanings were grouped, and categories and themes were defined. Then, the text was transferred into MAXQDA software, overarching themes were determined on the basis of the functions of education (individual, social, economic, political), and frequencies of the categories were considered to

reveal descriptively what kind of framework the core text provides concerning the functions of education. This analysis process enabled an in-depth evaluation of the theoretical framework of CTEM.

In the second phase of the analysis, the data obtained from the interviews were imported into MAXQDA 2024 software, read holistically, and meaningful statements were identified through open coding. Throughout the coding process, the data were systematically compared, and similarities and differences between teacher views and the findings of the core text analysis were identified. In addition, the constant comparison method was applied to ensure integrity and consistency among the data; data obtained from teacher views were interpreted together with the findings of the core text analysis. In this process, how and to what extent the functions of education (individual, social, economic, political) were addressed within the context of the *Century of Türkiye Education Model*, in what ways they were supported, and the strengths and weaknesses in practice were examined in detail. After the completion of the analysis process, the results obtained were presented in the form of graphs and comparative tables, and the data under each theme were supported with frequencies and direct quotations.

FINDINGS

In the study, comparative charts were created based on teachers' responses to whether the CTEM fulfills the individual, social, economic, and political functions of education, as evaluated through both document analysis and teacher interviews. In addition, Core categories were identified according to statements that serve these educational functions and presented in tabular form.

Finding 1: The Extent to Which CTEM Fulfills the Individual Function of Education

This section examines the extent to which CTEM fulfills the individual function of education, based on teacher perspectives.

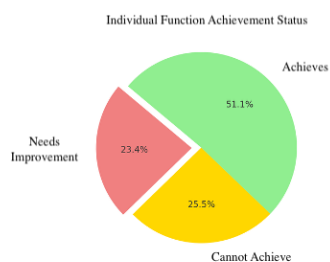


Figure 1. Extent to Which CTEM Fulfills the Individual Function of Education

The figure evaluates the extent to which the *Century of Türkiye Education Model* (CTEM) fulfills the individual functions of education. According to the data, 51.1% of participants stated that the model successfully fulfills these functions, while 25.5% indicated that it does not. Additionally, 23.4% believed that improvements are needed. These results suggest that there is generally a positive perception of the model's effectiveness in addressing individual functions; however, a significant portion of participants perceive notable shortcomings. In particular, the 23.4% who emphasized the need for improvement highlight the necessity for more concrete actions to enhance the model's implementation in terms of individual educational functions.

Table 2. Individual Function of Education: Comparison of the CTEM Core Text and Teacher Perspectives

Theme	CTEM Core Text	n	Teacher Perspectives	n
Individual Development and Personality Formation	Moral and Ethical Competencies	13	Educating individuals with moral values	2
			Supporting affective skills	2
	Emotional and Social Competencies	84	Fostering aesthetic sensitivity	1
			Developing responsibility skills	1
			Encouraging self-expression	1
			Cultivating appreciation for the arts	1
			Fostering cooperation skills	1
			Encouraging curiosity	1
			Building self-confidence	7
			Enhancing self-efficacy	1
Cognitive and Intellectual Abilities	Physical and Bodily Competencies	17		
	Mental and Creative Competencies	37	Fostering open-mindedness	1
			Encouraging creative and innovative thinking	1
			Developing original and innovative thinking skills	2

Theme	CTEM Core Text	n	Teacher Perspectives	n
Teaching and Learning Environment Management	Information Processing and Critical Thinking Skills	46	Teaching subject-specific skills	3
			Educating individuals with analytical thinking	1
			Cultivating idea generation and problem-solving	1
			Fostering inquiry-based learning	1
			Supporting decision-making	1
			Supporting problem-solving skills	1
			Enhancing critical thinking	5
			Supporting reasoning	2
			Supporting inferencing	1
			Developing inquiry skills	1
	Supporting Teaching Processes	63	Reducing learning difficulties	3
			Preparing students for life	1
			Supporting students' strengths (enrichment)	2
			Considering individual differences	1
			Allowing self-paced progress	1
			Addressing students' deficiencies	2
			Promoting student engagement	3
			Supporting individual activities	1
			Process-oriented assessment	4
			Developing individual talents	3
			Enabling self-discovery and self-awareness	1
Holistic Approach and Lifelong Learning Challenges Encountered in Practice	Adaptation to Contemporary Requirements	34		
	Holistic Individual Development	15		
			Leading to imitation rather than self-reflection	1
			Failing to consider differences in student ability levels	1
			Excessive repetition in certain activities	1
			Lack of sufficient pilot implementation	1
			Inadequate student comprehension and interpretation	2
			Failure to account for overcrowded classrooms	2
			Ignoring time constraints	1
			Insufficient support for academic development	1
			Overlooking regional disparities	2
			Neglecting physical infrastructure conditions	2

While CTEM offers a comprehensive framework targeting the holistic development of the individual, teacher perspectives provide critical insights into how this framework translates into practice. The themes derived from teacher statements enable a multidimensional evaluation of the model's effectiveness in fulfilling the individual function of education.

Within the scope of Individual Development and Personality Formation, many teachers emphasized that students are given opportunities to explore their own potential. For instance, Participant T1 stated, *"The model aims for self-realization by taking individual differences into account,"* highlighting this aspect. This aligns with the emphasis in the Core Text, which underlines *"allowing each student to know and discover themselves."*

In terms of Emotional and Social Competencies, some teachers noted that the model supports self-confidence, responsibility, and communication skills. For example, Participant T19 remarked, *"I find it very valuable in how it supports children in discovering their own characteristics."* This perspective parallels the Core Text's statement: *"Social-emotional learning competencies encompass students' ability to recognize emotions, empathize with others, take responsibility, make healthy decisions, and build positive relationships. These competencies are structured as a core part of all courses within CTEM."*

In the context of Cognitive and Intellectual Abilities, a significant number of teachers believed that the model fosters skills such as critical thinking and problem-solving. Participant T12 supported this by stating, *"It aims to unlock the individual potential of students."* These abilities are echoed in the Core Text as *"generations who think critically, solve problems, and make decisions."*

Regarding Teaching and Learning Environment Management, it was noted that the model supports individual learning pace and differences. However, Participant T27 noted, *"Class sizes are large, which makes it difficult to support individual*

development,” pointing to some limitations encountered in practice. This contrasts with the theoretical emphasis in the Core Text on *“flexible learning environments and process-based assessment.”*

Implementation Challenges emerged as one of the most prominent issues in teacher perspectives. Participant T33 expressed a disconnect between the model’s ideal framework and its practical reflection by stating, *“The aspects that support individual development are important, but they are not felt in the field.”* This disconnect aligns with the shortcomings identified by teachers in areas such as guidance, class size, content alignment, and regional disparities.

While teacher opinions generally reflect satisfaction with the individual function of the Maarif Model, there is a clear call for improvement at the implementation level—particularly regarding content, teacher training, and physical conditions. This provides valuable insights into how theoretical goals correspond to realities in the field and underscores the need to reassess implementation processes.

In general, the Core Text presents a comprehensive vision for cultivating the ideal individual. However, teacher perspectives address how this vision plays out in practice and highlight the challenges encountered. While the theoretical framework of the Core Text embodies a holistic approach aiming at both individual and social development, teacher feedback offers critical clues for understanding the model’s practical implications and overcoming obstacles. The relationship between theoretical objectives and practical needs is thus a key factor that must be considered at both the planning and implementation stages of education.

Finding 2. The Extent to Which CTEM Fulfills the Social Function of Education

This section examines the extent to which CTEM fulfills the social function of education, based on teacher perspectives.

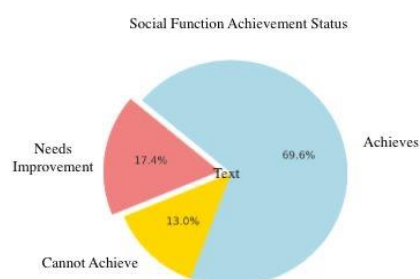


Figure 2. Effectiveness of CTEM in Fulfilling the Social Function of Education

The figure evaluates the extent to which CTEM fulfills the social functions of education. According to the data, 69.6% of participants stated that the model successfully fulfills social functions, while 13.0% indicated that these functions are not being achieved. Additionally, 17.4% of respondents believe that improvements are needed. These findings suggest that social functions represent one of the model’s strongest aspects and are largely perceived positively by teachers. However, the 17.4% who called for further development highlight that there are still shortcomings to be addressed in this area. This underscores the importance of ongoing evaluation and improvement efforts to enhance the model’s effectiveness in fulfilling its social functions.

Table 3. Social Function of Education: Comparison Between the Core Text of CTEM and Teacher Perspectives

Theme	Core Text	n	Teacher Perspectives	n
Social Cohesion and Solidarity	Social Responsibility and Ethical Awareness	44	Educating individuals who are aware of virtues	4
			Educating honest and moral citizens	1
			Educating individuals with ethical values	1
			Educating individuals with moral values	3
			Educating responsible individuals	6
			Educating contented individuals	1
	Common Culture and Value Consciousness	10	Educating individuals committed to traditions	1
			Educating individuals aware of national and societal values	14
	Strengthening Family and Social Bonds	61		
	Environmental and Aesthetic Awareness	22		
Socio-Emotional Learning and Communication	Social and Communicative Skills	57	Supporting socialization	5
			Developing social awareness	3
			Strengthening cooperation and solidarity	8
			Supporting social integration	3

Theme	Core Text	n	Teacher Perspectives	n
Historical and National Consciousness	Self-Actualization and Social Participation	4		
	Historical Consciousness and Empathy	6		
	National, Spiritual, and Cultural Identity	16	Educating individuals aware of their history	1
Holistic Development			Educating individuals aware of their culture	8
			Developing spiritual awareness	3
Holistic Development	Holistic Development and Wisdom	1		
	Holistic Quality of Life	25		
Challenges in Implementation			Negative influence of social media	1
			Continuation of lack of rules and sanctions	1
			Imposition of educational policies	1
			Lack of school-family-community integration	2

CTEM conceptualizes the social function of education not merely as citizenship education, but as a holistic construct encompassing cultural continuity, social participation, social responsibility, and the development of empathy. Teacher perspectives provide critical insights into the practical reflections of this multidimensional understanding of education's social function. Based on the analyses, five key themes have emerged from teacher responses: Cultural Heritage and Moral Values; Social Participation and Responsibility; Cooperation and Social Sensitivity; Socio-Emotional Development and Parental Involvement; Social Awareness and Realism.

Cultural Heritage and Moral Values: A majority of teachers positively evaluated the model's emphasis on sustaining and transmitting the fundamental values of society to future generations. For instance, Participant 5 emphasized this aspect by stating, *"The model supports moral development by transmitting the society's cultural values to students."* However, some teachers noted that this transmission remains largely theoretical and lacks sufficient practical implementation. Teacher 26, for example, remarked, *"The content for transmitting social values is inadequate; it needs to be more practice oriented."* The Core Text of the model aligns with this intent by defining its aim as *"transmitting the society's world of values to the new generation by harmonizing it with the requirements of the era,"* and foresees value education to be integrated across all subject areas.

Social Participation and Responsibility: CTEM aims to foster integration of students into society, cultivate sensitivity to social issues, and encourage active participation in social responsibility projects. Teacher 18 positively assessed this practical orientation by stating, *"It prepares students for social life through social responsibility projects."* In contrast, Teacher 7 noted, *"The program is very idealistic, but the field conditions do not support it,"* and T33 stated, *"Classroom settings are not sufficient for the development of social skills and responsibility,"* thus pointing to structural inadequacies in implementing these goals. These remarks suggest the presence of systemic barriers in realizing objectives such as *"community service practices, local projects, and empathy with disadvantaged groups,"* as outlined in the Core Text.

Cooperation and Social Sensitivity: One of the key components of the social function—coexistence and collaboration—was seen by several teachers as a strength of the model. For example, Teacher 21 stated, *"It foregrounds values like cooperation, solidarity, and mutual support."* The Core Text conceptualizes these competencies as *"developing social inclinations toward understanding different segments of society,"* and encourages group work, collective activities, and local culture-based initiatives.

Socio-Emotional Development and Parental Involvement: Another complementary aspect of the model's social function involves emotional competencies and school-family collaboration. Teacher 14 highlighted this limitation by stating, *"It is strong in terms of socio-emotional development, but parental involvement is lacking in practice."* While the Core Text indirectly acknowledges this shortcoming by defining families as *"a natural support system for children's healthier integration into social life,"* teacher statements indicate that this definition is not adequately reflected in practice.

Social Awareness and Realism: This theme reflects on the model's aim to raise awareness about current social events and issues. Teacher 30 critically stated, *"It fails to create awareness of societal problems; it remains too abstract."* This critique reflects a gap between the model's idealized goals and their tangible realization in practice. Yet, the Core Text explicitly aims for *"raising students who are sensitive to social issues, who take social responsibility, and who develop empathy."*

In summary, while teachers generally find the theoretical framework of CTEM supportive of education's social function, they also highlight several implementation shortcomings—especially in terms of content adequacy, physical infrastructure, teacher training, and school-family cooperation. It may be concluded that in order to concretize the model's goals related to social responsibility, value transmission, and social participation, more practice-oriented policies and stronger structural supports are needed.

Ultimately, this comparison between the Core Text and teacher perspectives reveals that policymakers offer a broad vision for achieving societal goals, whereas teachers focus on daily educational practices and individual student outcomes. Aligning these two perspectives is crucial for the feasibility and success of educational policies. Balancing the individual and social functions of education will serve as a bridge to close the gap between theoretical goals and practical realities.

Finding 3. The Extent to Which CTEM Fulfills the Economic Function of Education

This section examines the extent to which the Century of Türkiye Education Model fulfills the economic function of education, based on teacher perspectives.

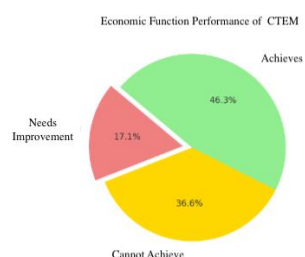


Figure 3. The Extent to Which CTEM Fulfills the Economic Function of Education

The graph evaluates the extent to which CTEM fulfills the economic functions of education. According to the data, 46.3% of participants stated that the model fulfills the economic functions, while 36.6% indicated that these functions are not being realized. The proportion of those who believe improvements are necessary stands at 17.1%. These results suggest that the economic functions are perceived as weaker compared to the other functions of the model. This relatively low success perception regarding economic functions highlights the need for improvements in areas such as vocational education, alignment with the labor market, and the enhancement of individuals' economic competencies. Notably, the combined percentage of "needs improvement" and "not fulfilled" responses (53.7%) points to the necessity of taking more concrete steps in this area.

Table 4. Economic Function of Education: Comparison Between the CTEM Core Text and Teacher Perceptions

Theme	Core Text Categories	n	Teacher Perspectives	n
Economic Competence and Economic Behaviors	Financial Awareness and Literacy	53	Not Developing Financial Literacy Skills	1
			Supporting Effective Time Management	2
Education and Resource Management	Resource Management and Saving	13	Raising Thrifty Individuals	1
			Increasing Paper Waste	1
			Raising Skill-Based Individuals	1
			Raising Productive Individuals	2
			Enabling Individuals to Realize Their Potential	3
			Multi-dimensional Development of Individuals	1
			Strengthening the Alignment Between the Education System and the Labor Market	1
			Supporting Vocational and Technical Education	2
			Raising Content Individuals (Satisfactory, Modest)	1
			Guiding Towards the Right Career Choice	3
Economic Contribution and Development	Economic Value and Production Awareness	1	Raising Individuals with Vocational Competence	2
			Teaching New Job Fields and Sectors	1
			Developing Technologically and Contemporary Skills	3
			Raising Productive Individuals Without Financial Necessity	1
			Raising Individuals with Qualified Workforce Skills	5
			Raising Economically Strong Individuals	2
			Raising Economically Conscious Individuals Regarding Efficiency	5
			Education Being Compulsory	
			Lack of Establishment of Vocational Middle Schools	
			Insufficient Economic Conditions of Students	
Economic Challenges and Implementation Areas			Not Preventing Unemployment	
			Economics Knowledge Remaining Abstract at the Primary Education Level	14
			Not Training Intermediate Staff	
			Insufficient Student Recruitment and Guidance in Artistic and Sports Fields	
			Emphasis Only on Quantitative (STEM) Fields	

Theme	Core Text Categories	n	Teacher Perspectives	n
			Prioritizing Individual Success Over Social Development	
			Schools Not Being Economically Supported by the Ministry	

CTEM adopts an approach that aims to equip individuals with competencies in areas such as economic awareness, financial literacy, and support for domestic production. Teachers' views provide a multidimensional opportunity to assess how these goals are reflected in practice. The themes shaped by their opinions reveal both the strengths of the model in fulfilling its economic function and the challenges encountered during implementation.

Financial Literacy and Economic Awareness: While most teachers viewed CTEM's goals related to financial literacy as positive in terms of content, they noted that these goals are not fully reflected in classroom practices. For example, P19 stated, "*Children should be taught money management from an early age—the model is good, but implementation is lacking,*" while P23 remarked, "*Financial literacy topics are present in the curriculum, but it's unclear how to implement them in practice.*" These statements highlight a gap between theoretical aims and actual application. Although the Core Text outlines objectives such as "*imparting basic knowledge on income, expenditure, saving, investment, borrowing, and financial technologies,*" classroom implementation needs to be more robustly structured.

Resource Management and Saving Awareness: Some teachers noted that while the model promotes common values on the conscious and planned use of resources, there is uncertainty about how these ideas translate into students' daily lives. P11 commented, "*Topics like saving and effective resource use are covered in lessons, but they need to be connected to real life.*" The Core Text defines this theme as "*the efficient and planned use of resources like time, energy, and money, avoiding waste.*" However, teacher feedback suggests that these concepts should be concretized, tailored to age levels, and supported by activities that foster behavioral transformation.

Domestic Production and National Economic Consciousness: CTEM's objective to support national production and emphasize domestic values was appreciated by some teachers, who also emphasized the need for practical applications to convey this to students. For example, P8 stated, "*Incorporating the country's production power, domestic products, and values like frugality into education is important,*" while P2 added, "*The model provides content that supports the national economy, but it's difficult to engage students—pedagogical approaches need diversification.*" Although the Core Text aims to "*instill consumer behavior that supports the national economy and raise awareness of domestic production,*" teachers believe that to turn this awareness into behavior, the model should be designed in a more interactive and case-based manner. Teacher feedback shows that CTEM includes strong theoretical goals regarding its economic function, but that structural and pedagogical difficulties arise in implementing these goals in the classroom. While theoretical attainments in financial literacy, resource management, and national economic awareness are supported by teachers, they emphasize the need to concretize the content, adapt it to student age groups, and diversify the tools used in practice. This indicates that in order for the model to effectively fulfill its economic function, adjustments are needed in content design, teacher competence, and learning environments.

In conclusion, the Maarif Model's text highlights micro-level strategies aimed at developing individuals' economic competencies, whereas teacher perspectives draw attention to broader economic functions and practical challenges. For educational policies to enhance alignment with labor markets, support individuals in realizing their economic potential, and strengthen the role of education in social development, both individual and societal economic goals must be addressed in a balanced way. This analysis underlines the importance of shaping policies in harmony with on-the-ground realities of implementation.

Finding 4: CTEM's Fulfillment of the Political Function of Education

In this section, the extent to which the Türkiye Century Education Model fulfills the political function of education has been examined based on teachers' perspectives.

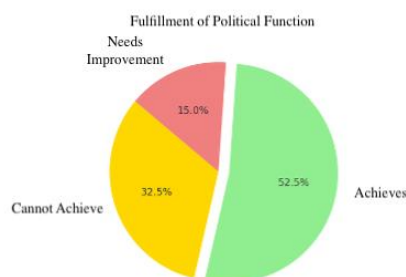


Figure 4. The Extent to Which CMET Fulfills Its Political Function

The graph evaluates the extent to which CTEM fulfills its political functions. According to the data, 52.5% of the participants stated that CTEM fulfills its political functions, while 32.5% believed these functions are not being achieved. The proportion of

those who think improvements are needed is 15.0%. These results indicate that CTEM has achieved a certain level of success in fulfilling its political functions. However, it is also evident that some shortcomings need to be addressed in order to fully realize these functions. In particular, it appears that the implementation of goals such as fostering democratic attitudes, developing a sense of citizenship, and raising individuals committed to national values should be made more concrete. The 15.0% who emphasized the need for improvement suggest that stronger support mechanisms are required in the implementation of this function.

Table 5. Political Function of Education: Comparison Between CTEM Core Text and Teachers' Perspectives

Category	Statements from the Core Text	n	Teacher Perspectives	n
Political Participation and Leadership	Empowering educational approach for social participation and equality	5	Raising Political Awareness	3
			Developing a Participatory Management Approach	1
			Facilitating the Development of Leadership Qualities	1
			Ensuring Active Participation of Individuals in Political Processes	6
National Values and Global Approaches	Fostering critical sociological thinking skills	2	Developing Critical Thinking	1
	Taking measures to eliminate disadvantages such as faith, identity, and socioeconomic status	1	Aiming to Balance Local Values and Global Requirements	2
			Instilling Concepts of Equality and Justice	2
	Educating Individuals Who Know Their Citizenship Rights and Responsibilities	18	Ensuring Respect and Tolerance for Different Opinions and Cultures	3
			Emphasizing Values of Homeland, Flag, and Nation	1
			Aligning with Global Political Values	0
			Instilling Citizenship Rights, Freedoms, and Responsibilities	3
	Citizenship Education Based on Civilizational Awareness, Equity, and State Responsibility	12	Fostering National Consciousness	1
			Raising Individuals Who Protect Their Nation, History, and Language within a Framework of Virtue and Value Action	1
			Educating Individuals Loyal to Atatürk's Principles and Reforms	2
Citizenship Awareness and Democracy Education	Ensuring a Learning Process Grounded in Rights and Personal Development	2	Educating Individuals Who Know Their Rights	1
			Raising Individuals Who Will Contribute to the Advancement of Their Country	1
			Developing Perspectives Suitable for Rapidly Changing Political Functions	1
			Fostering a Sense of Belonging to the State and Nation in Individuals	1
			Emphasizing Democracy, Human Rights, and the Rule of Law	1
			Raising Individuals Sensitive to Social Issues	2
			Developing Democratic Awareness	12
	An Inquiry- and Interdisciplinary-Thinking-Based Educational Model	10		
Systemic and Implementation Challenges			Content Issues Regarding Interdisciplinary Harmony and the Place of Atatürk's Thought in Education	2
			Insufficiency in Terms of Content and Implementation	3
			No Difference from Old Programs, Repetition	4
			Students' Digital Addictions and Negative Influences	1
			Ideological Content in Social Sciences Courses	1
			Insufficient Recognition of Different Cultural Identities and Social Inequality	3
			Differences in Education Quality Between Regions	1
			Providing Only Theoretical Knowledge	1
			Lack of Support for Applied and Experiential Learning Environments	1
			The Political and Ideological Thought Structure of the Student's Family Environment	2
			Failure to Foster National Consciousness Through Education	1
			Lack of Trust in the System	2

Political Participation and Leadership: The theme of political participation and leadership is directly related to the mission of CTEM to prepare individuals for active citizenship roles. Teacher 13 shared the view, *"Individuals actively participate in these processes,"* while T6 pointed out, *"Leadership qualities emerge,"* emphasizing that students' leadership tendencies are supported. These statements align with the theoretical text of CTEM, which highlights that *"Contents aimed at raising the society's leading individuals are included"* and *"Social studies, history, and revolution history courses aim to raise political awareness."* Teacher opinions largely support CTEM's focus on political consciousness and leadership.

National Values and Global Approaches: The emphasis that CTEM should both strengthen national identity and raise individuals who adapt to the global world was distinctly present in teacher perspectives. T15 stated, *"I find it sufficient in terms of imparting national values, creating national consciousness, and even raising leaders,"* while T8 noted, *"The Educational Model is structured both to protect local identity and to prepare students for a global world,"* pointing to a dual educational goal. This approach corresponds to the Core Text's emphasis that *"Content strengthening national identity is prioritized"* and *"Developing tolerance towards different cultures and communities ensures social integration."*

Citizenship Awareness and Democracy Education: Citizenship awareness and democracy education are among the fundamental political functions of CTEM. Teacher 4 said, *"It effectively develops citizenship awareness,"* while T9 shared, *"Democracy education aims to cultivate citizenship awareness and political participation."* The Core Text includes statements such as *"The goal is to develop citizenship awareness and raise individuals committed to democratic values"* and *"Concepts like democracy, human rights, and rule of law are addressed."* There is strong parallelism between the perspectives and the theoretical framework.

Systemic and Implementation Challenges: Challenges encountered in field implementations of CTEM were frequently expressed in teacher perspectives. T10 commented, *"Although it includes many aspects in terms of political function, it falls short of fully realizing them,"* and T25 stated, *"The social sciences courses in education are already completely ideological and their functionality is at the highest level."* These criticisms may contrast with the theoretical neutrality emphasized in the text: *"The model is not prepared with the aim of promoting a particular political ideology."* Teachers' observations during implementation reveal discrepancies between CTEM's ideal goals and field conditions.

In conclusion, while CTEM Core Text offers a theoretical framework regarding the political functions of education, teacher views assess the practical reflections of this framework from a more pragmatic perspective. The education system should not only teach individuals political values but also provide experiential environments in which they can apply these values in everyday life. In this context, it is of great importance that education policies support inclusive, egalitarian, and democratic values at both the theoretical and practical levels.

DISCUSSION, CONCLUSION AND SUGGESTIONS

This study aimed to evaluate the CTEM in terms of its individual, social, economic, and political functions of education based on teachers' perspectives. The findings revealed that CTEM addresses these four core functions in a holistic manner; however, individual development appears to be at the center of CTEM, which seeks to foster individuals' intellectual, emotional, moral, and aesthetic dimensions, supporting them comprehensively through elements such as critical thinking, creative problem-solving, and value-based approaches. Although the frequency distributions and graphical analyses did not show sharp quantitative differences among the four dimensions, both the Core Text and teacher perspectives consistently placed stronger emphasis on individual development. Teachers frequently highlighted themes such as personalized learning, talent discovery, and socio-emotional support, while the official document repeatedly foregrounded intellectual, emotional, moral, and aesthetic aspects. Thus, the interpretation of individual development as central rests not on numerical dominance but on the qualitative intensity and salience of references across data sources.

According to teachers' perspectives, CTEM adopts an approach that prioritizes individual development. While frequency counts were calculated during coding, most of them were very low (one or two instances), which limits their suitability for comparative interpretation. Therefore, frequency data were used only descriptively to illustrate the presence of themes, while the interpretation relied primarily on thematic patterns and direct teacher quotations that provide richer insights. One teacher emphasized this by stating, *"The model allows each student to progress at their own pace,"* highlighting the opportunity for personalized learning. This approach aligns with findings that demonstrate the positive impact of individualized learning models on student achievement (Tomlinson, 2001; Aypay, 2020). Similarly, the Core Text supports this approach with the statement, *"By allowing each student to recognize and explore their identity... free learning environments have been expanded."*

Addressing sensitivity to individual needs, another teacher remarked, *"It's a model that appeals to every type of student."* This aligns with the core text's assertion: *"Knowledge, skills, dispositions, and values are addressed by being reinforced with talents, interests, needs, and individual differences."* Regarding student self-expression, one teacher noted, *"It enables students to express their emotions freely,"* which corresponds with the core text's statement: *"Language is considered a key tool for students to make sense of themselves and others."*

Regarding critical thinking, one teacher stated, *"It instills critical thinking."* This emphasis aligns with research highlighting the pivotal role of critical thinking—one of the key 21st-century skills—in fostering holistic individual development (Argon & Selvi, 2011). CTEM Core Text supports this objective with the statement, *"Raising generations who think critically, solve problems, and make decisions is among the core principles of our education system."* Another teacher pointed out, *"Focusing on process evaluation increases the effectiveness of individual functions,"* referring to the model's process-oriented approach. The Core Text reinforces this by stating, *"Process- and situation-based assessment and evaluation approaches are included."*

One teacher remarked, *"I believe it is particularly effective in revealing students' talents,"* expressing a view that supports CTEM's emphasis on discovering individual potential. This perspective aligns with learner-centered curricula that prioritize students' interests and abilities (Varış, 1991; Güneş, 2020). The CTEM text reflects this approach with the statement, *"The model aims to help students discover their areas of interest and develop their talents."* Regarding social-emotional development, a teacher commented, *"Children participate in activities with a sense of self-confidence."* This observation corresponds with the CTEM statement, *"Social-emotional learning skills are an essential component of all courses."* The significance of these skills in the learning process is frequently emphasized in the literature (Durlak et al., 2011; Güleç & Alkış, 2022).

However, some teachers expressed concerns about CTEM's feasibility. One teacher stated, *"With the current system, it is not possible for students to discover themselves; the physical conditions are inadequate."* This view aligns with research emphasizing that implementer capacity and resource adequacy are critical for the success of educational reforms in the field (Fullan, 2007; Gündoğdu & Tekin, 2023). CTEM offers a solution-oriented approach by stating, *"Flexibility will be provided through lesson hours that can be allocated according to needs for each course."* Another teacher remarked, *"All children are equal, but regional differences have not been considered."* CTEM addresses this issue with the statement, *"Local and regional educational needs will be taken into account."*

Teacher views on the social function of CTEM revolve around themes such as social responsibility, the transmission of cultural values, cooperation, and social awareness. One teacher emphasized fostering social consciousness by stating, *"It provides the necessary opportunities for the student to realize they are part of society and become a beneficial individual to the community."* This is supported in the Core Text by the sentence, *"The goal is to raise individuals who act with social awareness."* In this context, education should not only focus on individual development but also instill the responsibility of producing social benefit (Durkheim, 2012; Hoşgörür & Taştan, 2012b). Highlighting the aspect that encourages socialization, one teacher commented, *"It raises students to be more social individuals."* The Core Text defines this aim as, *"The development of students' social-emotional skills is a fundamental component of educational programs."* Socialization is one of the most important educational processes that enable the individual to integrate with society (Tezcan, 1991).

The topic of social responsibility is highlighted by a teacher's view stating, *"The aspect of preparing students for social life through social responsibility projects is positive."* This aspect is explained in the Core Text with the statement, *"Community-beneficial activities are supported through the social responsibility program."* Social responsibility education is one of the fundamental tools that strengthen social cohesion (Özdemir & Atmaca, 2022). In the context of virtue and value transmission, a teacher noted, *"There is a strong emphasis on virtue and values in the program."* This directly aligns with CTEM's foundation, expressed as *"The Virtue-Value-Action framework forms the basis of the program."* Value education guides students in their social lives (Kıncal & Kartal, 2015). This aspect of CTEM, which centers on individual development, is consistent with Ülçay's (2024) work emphasizing the holistic development of students' intellectual, emotional, and moral capacities.

Regarding collaboration, one teacher stated, *"The Educational Model develops students' ability to collaborate through group work,"* while the Core Text includes the statement, *"It aims to raise individuals who collaborate and are open to communication."* Collaborative learning is an important tool for fostering empathy and understanding in social life (Johnson & Johnson, 2004). Concerning the transmission of cultural heritage, a teacher expressed, *"It is very important for passing cultural heritage to new generations."* This aspect is reflected in the core text by the sentence, *"The transmission of national culture to future generations is targeted."* Cultural continuity is one of the fundamental social functions of education (Sönmez, 1991).

Regarding social norms, one teacher commented, *"I think it is successful in understanding and perceiving the norms and cultural values of our society."* The Core Text addresses this goal with the statement, *"It aims to raise individuals with social consciousness and a sense of social responsibility."* This approach highlights the role of education in the sustainability of social order (Aybek, 2007). While one teacher shared the perspective, *"We struggle a lot with implementation because it is not functional,"* the core text responds to this criticism with, *"Flexibility will be provided according to local needs in the implementation of curricula."* In this context, the localization of education policies is crucial for the effective implementation of programs (Sahlberg, 2011).

Additionally, the teacher's perspective stating, *"An honest citizen is a moral individual. Therefore, an honest society is a moral society,"* aligns with the Core Text's statement, *"With the understanding that a moral individual shapes a moral society, individual development will influence the social structure."* Moral education is indispensable for the healthy development of social structures (Kohlberg, 1981).

Economic and political functions are less emphasized yet fundamental components of CTEM. According to teacher perspectives, CTEM aims to develop a qualified workforce, enhance productivity, foster a savings consciousness, and provide

vocational guidance within the scope of its economic functions. One teacher stated, *"The goal is to achieve quality production by training qualified personnel. If this is realized, it will make a significant contribution to the economy."* The core text emphasizes this objective with the sentence, *"Training qualified human resources, increasing productivity, and supporting economic development are fundamental goals."* A qualified workforce plays a strategic role in the transition to the knowledge economy (Castells, 2000; Tunali, 2019).

Regarding vocational guidance, one teacher stated, *"I find it positive that students are being prepared for their future professional roles, career choices, and vocational training."* The Core Text addresses this with the statement, *"Priority will be given to vocational and technical education to enable early workforce orientation."* Education-employment alignment is a prerequisite for the success of economic functions (Yıldız, 2006, Ünsal & Gökdaş, 2020). On the theme of productivity, a teacher emphasized, *"The Educational Model supports productivity by providing students the opportunity to engage in the process of creating a product through 'Production Workshop' activities."* This aspect of the model is described as, *"Production-focused skills will be supported through workshop applications."* The integration of education and production is essential for sustainable development (Ertürk, 1979).

Regarding savings awareness, one teacher stated, *"It distinguishes between saving and stinginess."* The Core Text supports this approach with the statement, *"Values of saving, responsibility, and sharing have been integrated into the curriculum."* Financial awareness, when developed at an early age, guides lifelong economic behaviors (Lusardi & Mitchell, 2014). Concerning financial literacy, a teacher remarked, *"Even if financial literacy is not present, courses aiming to foster this awareness should be incorporated into the curriculum or the model."* The common text addresses this by stating, *"Themes of financial literacy and economic awareness are included among the new elective courses."* This theme is directly related to individuals' economic citizenship skills in modern economies (Remund, 2010).

Regarding productivity, one teacher stated, *"I believe it will ensure productivity."* This perspective aligns with the Common Text's sentence, *"Individual production processes based on productivity are prioritized."* Learning environments focused on productivity also support entrepreneurial skills (Kuratko, 2005). From the perspective of economic equality, a teacher expressed, *"Ensuring equality for everyone is economically positive."* This aspect of CTEM is explained as, *"The principle of equal opportunity will guarantee everyone's right to education and economic participation."* Equality of opportunity in education is a fundamental aspect of social justice (Rawls, 1971).

However, some teachers indicated that CTEM struggles to fulfill this function at the implementation level. One teacher stated, *"The model is far from achieving its economic functions because even in education, the savings approach is the priority within the Ministry of National Education bureaucracy."* In response to such criticisms, the Core Text states, *"The sustainability of the system will be ensured through the effective and equitable use of resources."* System sustainability is only possible through fair distribution of resources (OECD, 2012).

Teacher perspectives on the political functions of CTEM have focused on themes such as citizenship awareness, the acquisition of democratic values, the development of leadership skills, and the formation of national identity. One teacher stated, *"It effectively develops citizenship awareness,"* highlighting the model's contribution to civic education. In the Core Text, this goal is defined as *"The development of citizenship awareness and the cultivation of individuals committed to democratic values are targeted."* Citizenship education is indispensable for the sustainability of democracy (Kerr, 1999).

In the context of democratic values, one teacher stated, *"By providing democracy education, the aim is to foster citizenship awareness and political participation."* This aligns with the statement, *"It is aimed to include concepts such as democracy, human rights, and the rule of law."* Democratic educational environments aim not only to impart knowledge but also to instill values (Gutmann, 1999). Regarding political participation, a teacher noted, *"Individuals actively participate in these processes,"* which parallels the statement, *"Courses such as social studies, history, and revolution history aim to raise political awareness."* Political literacy is a fundamental quality of the modern citizen (Galston, 2001).

Regarding leadership tendencies, one teacher stated, *"Leadership qualities are emerging."* The Core Text reflects this by stating, *"Contents aimed at training society's leader individuals are included."* Leadership in education increases an individual's capacity for social participation (Leithwood & Jantzi, 2005). Additionally, a teacher's perspective that *"Since the model supports critical thinking... it also supports participation in democratic processes"* directly relates to the statement, *"It is aimed to create an educational environment that supports critical thinking."* Critical thinking is a fundamental skill in the development of individuals within democratic societies (Brookfield, 1987).

Regarding national identity and values, one teacher stated, *"I find it sufficient in imparting national values, fostering a sense of nationhood, and even raising leaders."* This view is reflected in CTEM with the statement, *"Emphasis is placed on content that strengthens national identity."* Education on national identity enhances an individual's sense of social belonging (Smith, 1991). Additionally, a teacher's perspective that *"The Educational Model is structured to both preserve local identity and prepare students for a globalized world"* is supported by the sentence, *"Developing an understanding of tolerance toward different cultures and communities ensures social integration."* Education should cultivate both national and global citizenship awareness (Banks, 2008).

Some teachers, however, emphasized that CTEM does not fully fulfill its political functions. The statement, *“Although it includes many aspects in terms of political functions, it remains insufficient in fully implementing them,”* draws attention to gaps at the implementation level. In response to this critique, the official document states, *“The extent to which these functions are realized requires application- and outcome-oriented analysis.”* The effectiveness of citizenship education depends on practice-based teaching methods (Demirbolat, 2019). Additionally, a teacher’s remark, *“The social science courses in education are completely ideological anyway, and their functionality is at the highest level,”* reflects a perception regarding CTEM’s ideological nature. The official document counters this perception with the clarification, *“The model is not designed to promote any particular political ideology.”* Ideological neutrality in education is crucial for the development of critical thinking (Apple, 2004).

When evaluated in comparison with previous curriculum reforms, CTEM can also be seen as a response to certain gaps that remained unresolved. For example, the constructivist curriculum of 2005 and the skills-based revisions of 2017 emphasized learner-centered approaches and competencies, yet they were often criticized for limited attention to value-based education, insufficient integration of socio-emotional skills, and inadequate mechanisms for local adaptation. Teachers also reported that the earlier reforms did not sufficiently address the connection between education and labor market needs. CTEM, by incorporating the Virtue-Value-Action framework, explicit social-emotional learning objectives, and provisions for flexibility at the local level, attempts to fill these gaps. This shift suggests that the new model is not only a continuation but also a corrective step in Turkish curriculum policy. Nevertheless, whether these promises can be realized depends heavily on implementation capacity, resource allocation, and long-term sustainability. Therefore, a balanced policy analysis requires recognizing both the innovative dimensions of CTEM and the structural constraints that may hinder its effectiveness.

In conclusion, CTEM presents a structure that prioritizes the multifaceted development of the individual and aims to strengthen social values. However, according to teacher perspectives, more resources, implementation support, and local adaptations are needed for the effective application of its economic and political functions in practice. Thus, while CTEM draws attention with its ideal vision, practical challenges have also become evident. Compared to the literature, the most significant contribution of this study is that the CTEM has been evaluated through both its theoretical text and teacher experiences. In this respect, the study stands out uniquely from existing research.

Based on the research findings, the following suggestions are offered to enhance the effective implementation of CTEM:

1. **Consideration of Individual Differences:** To support the individual-centered structure of the CTEM, flexible materials and teaching strategies that address different learning styles, prioritize individual development, and consider learning speeds should be developed. Process-oriented assessment and project-based learning approaches should be widely adopted.
2. **Collaboration with Families and the Community:** To ensure that students internalize the social functions, school-family and school-community cooperation should be strengthened. This collaboration should go beyond mere communication with parents and be supported by activities where social values are learned through lived experiences.
3. **Strengthening Economic Functions:** More investment should be made in vocational and technical education; individuals should be equipped with skills suitable to contemporary needs, such as digital literacy, financial awareness, and productivity. The link between education and the labor market should be reinforced.
4. **Translating Political Functions into Practice:** Activities aimed at fostering democratic values should be increased; participatory management approaches, citizenship education, and awareness of rights and responsibilities should be integrated into classroom activities. School councils and project-based community participation initiatives should be supported to help students gain active citizenship roles.
5. **Improving Implementation Processes:** The feasibility of the model should be continuously evaluated based on teacher feedback. Flexible implementation guidelines appropriate to local conditions should be prepared by considering field obstacles such as the lack of pilot applications, regional and physical differences, class sizes, and time management.
6. **Teacher Training and Involvement:** To implement CTEM effectively, teachers should be continuously supported through in-service training, actively involved in implementations, and included in decision-making processes. Teachers should be encouraged to take on roles not only as implementers but also as developers.
7. **Education on Values and Identity:** Themes such as commitment to national, spiritual, and ethical values; historical awareness; and cultural identity should be supported in practice not only at the discourse level but also through experience-based pedagogical methods. Students should be equipped with historical empathy and critical thinking skills.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Statements of publication ethics

We hereby declare that the study has not unethical issues and that research and publication ethics have been observed carefully.

Examples of author contribution statements

First author conceived of the research idea and conducted the data collection and initial analysis. Second author supervised the research process, guided the theoretical framework, and reviewed the manuscript critically. Third author contributed to the interpretation of the findings and provided revisions to improve the final version of the manuscript. All authors discussed the results and approved the final manuscript.

Researchers' contribution rate

The study was conducted and reported with equal collaboration of the researchers.

Ethics Committee Approval Information

This study was approved by the Ethics Committee of Sivas Cumhuriyet University on 15.03.2024 with decision number 2024/128.

REFERENCES

- Adem, M. (1977). *Eğitim planlaması*. Ankara Üniversitesi Eğitim Fakültesi Yayınları.
- Akpınar, B., & Köksalan, B. (2024). Eğitimde maarif ve müfredat yenileme ihtiyacı: Türkiye Yüzyılı Maarif Modeli üzerinden teorik bir analiz. *Tarih Okulu Dergisi*, 17(68), 27–48.
- Apple, M. W. (2004). *Ideology and curriculum* (3rd ed.). New York: RoutledgeFalmer.
- Argon, T., & Selvi, Ç. (2011). Öğretmen adaylarının eleştirel düşünme eğilimleri ve çatışma yönetimi stilleri. *2nd International Conference on New Trends in Education and Their Implications*.
- Arslankara, V. B., & Arslankara, E. (2024). Türkiye Yüzyılı Maarif Modeli'nin felsefi temelleri: Ontolojik, epistemolojik ve aksiyolojik bakış açılarından bir değerlendirme. *İstanbul Eğitim Dergisi*, 1(1), 121-145. <https://doi.org/10.71270/istanbulegitim.istj.1557889>
- Aybek, B. (2007). Eğitimin toplumsal işlevi ve değer eğitimi. *Değerler Eğitimi Dergisi*, 5(13), 45–66.
- Aypay, A. (2020). Güncel eğilimlerden hareketle gelecek eğitim politikalarına yönelik bir değerlendirme. *İnsan ve İnsan Dergisi*, 7(24), 1-20.
- Banaz, E. (2024). 2024 Türkiye Yüzyılı Maarif Modeli ortaokul türkçe dersi öğretim programı'nın dijital okuryazarlık açısından incelenmesi. *Kuram ve Uygulamada Sosyal Bilimler Dergisi*, 8(1), 279-290. <https://doi.org/10.48066/kusob.1483671>
- Banks, J. A. (2008). Diversity, group identity, and citizenship education in a global age. *Educational Researcher*, 37(3), 129-139.
- Becker, G. S. (1993). *Human capital: A theoretical and empirical analysis, with special reference to education* (3rd ed.). University of Chicago Press.
- Berk, Ş. & Özer, B. (2024). "Türkiye Yüzyılı" Maarif Modeli (TYMM) öğretim programlarına ilişkin düşünceler. *Journal of History School*, 73, 3241-3256.
- Brookfield, S. D. (1987). *Developing critical thinkers: Challenging adults to explore alternative ways of thinking and acting*. San Francisco: Jossey-Bass.
- Bowles, S., & Gintis, H. (1976). *Schooling in capitalist America: Educational reform and the contradictions of economic life*. Basic Books.
- Castells, M. (2000). *The rise of the network society* (2nd ed.). Oxford: Blackwell Publishers.
- Creswell, J. W. (2016). Araştırma deseni: Nitel, nicel ve karma yöntem yaklaşımları (M. Bütün & S. B. Demir, Çev.). *Eğiten Kitap*.
- Demirbolat, A. O. (2019). Vatandaşlık eğitimi uygulamaları: Eleştirel bir değerlendirme. *Eğitim ve Toplum Araştırmaları Dergisi*, 6(2), 150–170.
- Doğan, S. (2015). Eğitimin işlevleri. In C. T. Uğurlu (Ed.), *Eğitim Bilimine Giriş* (ss. 39–53). Eğiten Kitap.
- Doğan, S., & Aşlamacı, İ. (2025). Türkiye Yüzyılı Maarif Modelinin değer eğitimi açısından analizi. *Çukurova Araştırmaları Dergisi*. <https://doi.org/10.29228/cukar.80525>
- Durkheim, É. (2012). *Education and sociology* (W. S. F. Pickering, Trans.). Free Press. (Original work published 1922)
- Durlak, J. A., Weissberg, R. P., Dymnicki, A. B., Taylor, R. D., & Schellinger, K. B. (2011). *The impact of enhancing students' social and emotional learning: A meta-analysis of school-based universal interventions*. *Child Development*, 82(1), 405-432.
- Duyul, S., Duyul, Y., Kesman, M., & Kesman, M. (2025). Türkiye Yüzyılı Maarif Modeli konusunda öğretmen görüşleri. *Ulusal Eğitim Dergisi*, 5(3), 20262. <https://uleder.com/index.php/uleder/article/view/562>
- Efe, M. E. (2024). Türkiye yüzyılı maarif modeli: Eğitimde kültürel değerler ve dijitalleşme perspektifi. *Route Educational & Social Science Journal*, 11(6), 213-224.
- Ergün, M. (1994). *Eğitim sosyolojisine giriş* (5. Baskı). Ankara: Ocak Yayınları.
- Ertürk, S. (1979). *Eğitimde program geliştirme*. Ankara: Yelkentepe Yayınları.
- Freire, P. (2000). *Pedagogy of the oppressed* (30th anniversary ed., M. B. Ramos, Trans.). Continuum. (Original work published 1970)
- Fullan, M. (2007). *The new meaning of educational change* (4th ed.). New York: Teachers College Press.
- Galston, W. A. (2001). Political knowledge, political engagement, and civic education. *Annual Review of Political Science*, 4, 217-234.
- Gutmann, A. (1999). *Democratic education* (Revised ed.). Princeton, NJ: Princeton University Press.
- Gündoğdu, O., & Tekin, Y. (2023). Türkiye'de eğitim reformları ve uygulama sorunları. *Eğitim Reformu Girişimi Raporu*.
- Güleç, S., & Alkış, S. (2022). Sosyal-duygusal öğrenme becerilerinin ölçülmesi. *Millî Eğitim Dergisi*, 52(Özel Sayı), 1119-1136.

- Güneş, E. (2020). Program geliştirme modelleri üzerine bir inceleme. *Kuram ve Uygulamada Eğitim Yönetimi*, 13(13), 79–97. <https://doi.org/10.14527/630>
- Hevedanlı, H., İnce, H.H. , (2025). Türkiye Yüzyılı Maarif Modeli biyoloji dersi öğretim programının soft beceriler açısından incelenmesi. *The Journal of Academic Social Science*
- Hoşgörür, V., & Taştan, N. (2012a). Eğitimin işlevleri. In Ö. Demirel & Z. Kaya (Eds.), *Eğitim Bilimine Giriş* (ss. 283–305). Pegem Akademi. <https://doi.org/10.14527/9786053186366.10>
- Hoşgörür, T., & Taştan, C. (2012b). *Eğitim sosyolojisi*. Anı Yayıncılık.
- Johnson, D. W., & Johnson, R. T. (2004). Assessing students in groups: Promoting group responsibility and individual accountability. Thousand Oaks, CA: Corwin Press.
- Karataş, İ. H. (2024). Türkiye Yüzyılı Maarif Modeli hakkında. *Alanyazın*, 5(1), 6-11.
- Kerr, D. (1999). Citizenship education: An international comparison. *International Review of Curriculum and Assessment Frameworks Archive*.
- Kıncal, R., & Kartal, O. (2015). Teorik olarak yapılandırılmış sosyal ve bireysel katılım faktörlerinin ampirik olarak sınanması. *Gazi Üniversitesi Gazi Eğitim Fakültesi Dergisi*, 35(2), 231-265.
- Koçer, H. A. (1974). *Türkiye’de modern eğitimin doğuşu ve gelişimi 1873–1923*. İstanbul: Talim Terbiye Dairesi Yayınları.
- Kohlberg, L. (1981). The philosophy of moral development: Moral stages and the idea of justice. San Francisco: Harper & Row.
- Köçer, N. (2025). Türkiye Yüzyılı Maarif Modeli Türkçe dersi öğretim programı ve 5. sınıf türkçe ders kitabının kültür okuryazarlığı bağlamında incelenmesi. *Uluslararası Türkçe Edebiyat Kültür Eğitim (TEKE) Dergisi*, 14(1), 161-171. <https://doi.org/10.7884/teke.1589388>
- Krippendorff, K. (2004). *Content analysis: An introduction to its methodology* (2nd ed.). Sage Publications.
- Kuratko, D. F. (2005). The emergence of entrepreneurship education: Development, trends, and challenges. *Entrepreneurship Theory and Practice*, 29(5), 577-597.
- Kuzu, O. & Göçer, V. & Akçay, A. O. (2024). Türkiye Yüzyılı Maarif Modeli kapsamında İlkokul Matematik Dersi Öğretim Programı’nın incelenmesi. *RumeliDE Dil ve Edebiyat Araştırmaları Dergisi*, (41), 640-667. DOI: <https://10.5281/zenodo.13337757>
- Leithwood, K., & Jantzi, D. (2005). *A review of transformational school leadership research 1996–2005*. Leadership and Policy in Schools, 4(3), 177-199.
- Lincoln, Y. S., ve Guba, E. G. (1985). *Naturalistic Inquiry*. Newbury Park, CA: Sage.
- Lusardi, A., & Mitchell, O. S. (2014). The economic importance of financial literacy: Theory and evidence. *Journal of Economic Literature*, 52(1), 5-44.
- Millî Eğitim Bakanlığı. (2024). *Türkiye Yüzyılı Maarif Modeli Ortak Metni*. <https://tymm.meb.gov.tr/ortak-metin>
- OECD. (2012). *Education at a Glance 2012: OECD Indicators*. Paris: OECD Publishing.
- OECD. (2019). *OECD future of education and skills 2030: OECD learning compass 2030*. OECD Publishing. <https://www.oecd.org/education/2030-project/>
- OECD.(2020). *Education policy outlook 2020: Shaping responsive and resilient education in a changing world*. OECD Publishing. <https://doi.org/10.1787/3f435615-en>
- Özdemir, H., & Atmaca, T. (2022). Sosyal sorumluluk çalışmalarının öğrencilerin sosyal gelişimlerine etkisinin öğretmenlerin görüşlerine göre incelenmesi. *Uluslararası Liderlik Eğitimi Dergisi*, 6(2), 69-91.
- Patton, M. Q. (2014). Nitel araştırma ve değerlendirme yöntemleri (3. Baskıdan Çeviri). Çeviri Editörleri: Mesut Bütün, Selçuk Beşir Demir). Ankara: Pegem Akademi.
- Rawls, J. (1971). *A theory of justice*. Cambridge, MA: Harvard University Press.
- Remund, D. L. (2010). Financial literacy explicated: The case for a clearer definition in an increasingly complex economy. *Journal of Consumer Affairs*, 44(2), 276-295.
- Sahlberg, P. (2011). *Finish lessons: What can the world learn from educational change in Finland?* New York: Teachers College Press.
- Sezer, B. B., & Dedeoğlu, H. (2024). Maarif Modeli İlkokul Öğretim Programı’nda okuryazarlık becerileri: Finansal, görsel ve eleştirel. *Türk Eğitim Bilimleri Dergisi*, 22(3), 2007-2027. <https://doi.org/10.37217/tebd.1511168>
- Sönmez, V. (1991). *Eğitim felsefesi*. Ankara: Adım Yayıncılık.
- Smith, A. D. (1991). *National identity*. Reno: University of Nevada Press.
- Şişman, M. (2007). *Eğitim bilimine giriş*. Ankara: Pegem Akademi Yayıncılık.
- Tezcan, M. (1981). *Eğitim sosyolojisine giriş*. Ankara Üniversitesi Eğitim Fakültesi Yayınları.
- Tezcan, M. (1991). *Eğitim sosyolojisi*. Ankara: Anı Yayıncılık.
- Tomlinson, C. A. (2001). *How to differentiate instruction in mixed-ability classrooms (2nd ed.)*. Alexandria, VA: Association for Supervision and Curriculum Development (ASCD).
- Tunalı, İ. (2019). Türkiye’de nitelikli insan gücü yetiştirme politikaları. *İktisat ve Toplum Dergisi*, 9(3), 45-58.
- Tunç, S. (1969). *Türkiye’de eğitimin eşitsizliği*. Ankara Üniversitesi Eğitim Fakültesi Yayınları.
- Turan, Ş. & Nazıroğlu, B. (2024). Türkiye Yüzyılı Maarif Modeli’nin din öğretimi yaklaşımında birey ve toplumun ideolojik tahayyülü. *Rize İlahiyat Dergisi* 27, 15-33. <https://doi.org/10.32950/rid.1507279>.
- Türer, A. (2006). *Eğitim bilimine giriş*. Sivas: Dilek Matbaası.
- Uygun, K., & Akgül, G. (2024). Sosyal bilgiler öğretmenlerinin Türkiye Yüzyılı Maarif Modeline ilişkin görüşleri. *Dumlupınar Üniversitesi Eğitim Bilimleri Enstitüsü Dergisi*, 8(2), 81-102.

- Ülçay, O. (2024). Türkiye Yüzyılı Maarif Modeli değerlendirmesi. *Ulusal Eğitim, Toplum ve Dünya Dergisi*, 1(2), 70–75. <https://doi.org/10.5281/zenodo.11097248>
- Ünsal, S., & Gökdaş, A. (2020). Eğitimin işlevlerinin gerçekleşme düzeyinin öğretmen algılarına göre incelenmesi. *Cumhuriyet Uluslararası Eğitim Dergisi*, 9(4), 1133–1157. <https://doi.org/10.30703/cije.675779>
- Üzümcü, M. Abanoz, S. (2024). Program tasarımı yaklaşımları bağlamında Türkiye Yüzyılı Maarif Modeli Din Kültürü Ve Ahlak Bilgisi Dersi Öğretim Programı. *Mevzu – Sosyal Bilimler Dergisi* 12 (September 2024), 795-826. <https://doi.org/10.56720/mevzu.1495250>.
- Varış, F. (1991). *Eğitim bilimine giriş*. Ankara: Ankara Üniversitesi Basımevi.
- Yıldırım, A., & Şimşek, H. (2022). *Sosyal bilimlerde nitel araştırma yöntemleri* (13. baskı). Seçkin Yayıncılık.
- Yıldırım, Y., & Çalışkan, A. (2024). Türkiye Yüzyılı Maarif Modeli'nin 21. Yüzyıl insan profili açısından değerlendirilmesi. *Elektronik Eğitim Bilimleri Dergisi*, 13(26), 204-220. <https://doi.org/10.55605/ejedus.1548121>
- Yıldız, S. (2006). *Eğitim ekonomisi*. Milli Eğitim Yayınları.
- Yurdakal, İ. H. (2024). Türkiye Yüzyılı Maarif Modeli: 2024 İlkokul Türkçe Dersi Öğretim Programı'nın (1, 2, 3 ve 4. Sınıflar) incelenmesi. *Temel Eğitim*, 6(24), 76-88.



| Research Article / Araştırma Makalesi |

The Relationship Between Turkish Origin Students' Intercultural Sensitivity and Foreign Language Learning Attitudes

Türk Soylu Öğrencilerin Kültürlerarası Duyarlılıkları ile Yabancı Dil Öğrenme Tutumları Arasındaki İlişki

Gülnur AYDIN¹, Nezir TEMUR²

Keywords

1. Intercultural sensitivity
2. Turkish origin students
3. Foreign language learning attitudes
4. Intercultural consciousness

Anahtar Kelimeler

1. Kültürlerarası duyarlılık
2. Türk soylu öğrenciler
3. Yabancı dil öğrenme tutumu
4. Kültürlerarası bilinç

Received/Başvuru Tarihi
08.01.2025

Accepted / Kabul Tarihi
10.10.2025

Abstract

Purpose: The aim of this research is to evaluate the relationship between Turkish origin students' intercultural sensitivity and their attitudes towards foreign language learning.

Design/Methodology/Approach: The research adopted a correlational research method. The participants of the study were 509 Turkish origin students studying at state universities in Azerbaijan, Kazakhstan, Kyrgyzstan and Uzbekistan. Data were collected using *Intercultural Sensitivity Scale* and the *Attitudes Towards Foreign Language Learning Scale*. The collected data were analyzed using the SPSS 26.0 statistical package program.

Findings: Based on the analyses conducted using the SPSS 26.0 statistical package program, it was identified that age and nationality were determinant demographic variables for the total scores of intercultural sensitivity, whereas gender, age, nationality, and the number of foreign languages known were determinant demographic variables for attitudes toward foreign language learning. The study found that as the levels of intercultural sensitivity increased, the attitudes towards foreign language learning developed positively, and the predictor that most influenced attitudes towards foreign language learning was "respect for cultural differences".

Highlights: Given the influence of intercultural sensitivity on language learning attitudes and, consequently, on individual motivation, it is crucial to re-examine the resources and methodologies required for its enhancement. Accordingly, intercultural education programs may be designed to foster and strengthen the intercultural sensitivity of students of Turkish origin.

Öz

Çalışmanın amacı: Bu araştırmanın amacı, Türk soylu öğrencilerin kültürlerarası duyarlılıkları ile yabancı dil öğrenmeye yönelik tutumları arasındaki ilişkiyi değerlendirmektir.

Materyal ve Yöntem: Araştırmada korelasyonel araştırma yöntemi benimsenmiştir. Katılımcılar, Azerbaycan, Kazakistan, Kırgızistan ve Özbekistan'daki devlet üniversitelerinde öğrenim gören toplam 509 Türk soylu öğrenciden oluşmaktadır. Veriler, *Kültürlerarası Duyarlılık Ölçeği* ve *Yabancı Dil Öğrenmeye Yönelik Tutum Ölçeği* aracılığıyla toplanmıştır. Elde edilen veriler, SPSS 26.0 istatistik paket programında analiz edilmiştir.

Bulgular: SPSS 26.0 istatistik paket programında yapılan analizler sonucunda kültürlerarası duyarlılık toplam puanları için yaş ve uyruğun; yabancı dil öğrenmeye yönelik tutumlar için cinsiyet, yaş, uyruk ve bilinen yabancı dil sayısının belirleyici demografik değişkenler olduğu tespit edilmiştir. Araştırmada kültürlerarası duyarlılık düzeyi arttıkça yabancı dil öğrenmeye yönelik tutumların olumlu yönde geliştiği, yabancı dil öğrenmeye yönelik tutumu en fazla etkileyen yordayıcının ise "kültürel farklılıklara saygı" olduğu belirlenmiştir.

Önemli Vurgular: Dil öğrenme tutumuna ve dolayısıyla da bireysel motivasyona etkisi düşünüldüğünde kültürlerarası duyarlılığın geliştirilmesi için gereksinim duyulan kaynak ve yöntemlerin gözden geçirilmesi büyük önem taşımaktadır. Bu doğrultuda, Türk soylu öğrencilerin kültürlerarası duyarlılığını geliştirmeyi ve güçlendirmeyi hedefleyen kültürlerarası eğitim programları tasarlanabilir.

¹ Corresponded Author, Ankara Hacı Bayram Veli University, Institute of Graduate Education, Department of Teaching Turkish as a Foreign Language, Ankara, Türkiye; <https://orcid.org/0000-0003-0490-9580>

² Gazi University, Gazi Faculty of Education, Department of Turkish Education, Ankara, Türkiye; <https://orcid.org/0000-0002-8052-1927>

INTRODUCTION

Globalisation has created a network that makes countries interdependent in many aspects such as communication, transport, economy, politics, education and has placed great importance on the ability to manage both national and international encounters/relationships. This ability, which focuses on the quality of contact with different cultures, refers to the ability to discover differences in interaction styles, world views, perspectives and behaviours, and to reflect these in one's own behaviour. This process, which Altmayer (2002) calls as learning *cultural patterns of interpretation*, makes the basis for communication and cooperation in today's world, where cultural diversity is commonplace, without the problem of adapting to the unusual. Therefore, the need to acquire abilities that require the restructuring of intercultural consciousness is increasing day by day.

The main function of language is to create a bond between individuals based on communication. Since the level of communicative competence will be decisive in the formation of this bond, what is included in the scope of communicative competence is an issue that needs to be taken into consideration. In the views put forward on this subject, communicative competence is explicitly and implicitly associated with knowledge of form and usage by emphasising its relationship with the social field. The process that starts with the individual's acquaintance with language exhibits a very comprehensive view that includes grammatical appropriateness focusing on structure, sociolinguistic appropriateness focusing on sociocultural context, strategic appropriateness focusing on usage and discourse (Bachman, 1990; Canale & Swain, 1980; Canale, 1983; Hymes, 1972). It is important to what extent successful communication in intercultural contact situations corresponds to these characteristics. These characteristics can also be considered as the basic indicators of a state of intercultural awareness that will build intercultural contact.

Intercultural consciousness goes beyond an understanding of diversity to include a commitment to ethical principles, patience, tolerance, flexibility and adaptability, and refers primarily to intercultural communicative competence to create a new understanding that supports the coexistence of cultural groups. Xin (2007) attributes the importance of communicative competence in today's world to its significance in the gradual globalisation and also points out that it has created a dynamic in language learning and teaching. This has led to changes in people's attitudes towards language learning in many parts of the world. In addition, culture, as one of the most important components of the communication process, has created paradigm shifts in language teaching approaches and it has often been stated by researchers that it should be an important part of planning the language teaching process (Aydin, 2021; Crawford-Lange & Lange, 1987).

It is stated that language learning improves cognitive and analytical abilities, increases awareness of one's own language and culture, strengthens global understanding, life skills, communication and relationships (Perić & Radić, 2021). With these characteristics, knowing a language (or languages) other than one's mother tongue is a symbol of power, equivalent to having many opportunities in different cultures. According to Kramsch (1998), culture is membership in a discourse community that shares a common social space, history and dreams. Knowledge of the background of the discourse, extra-linguistic indicators and social skills enable the language code to be analysed correctly. Thus, interpretation emerges and is constructed from social interactions (Kaufman, 2004; Nunan, 1992). Aidinlou and Kejal (2012) found that perceiving the sociocultural features of the target language contributes significantly to learning that language better, deciding where, how and when to use language structures, and saving learning time. The distance between the source language and the target language can be reduced by strengthening the cultural background. From this point of view, culture accompanying language learning can be seen as creating meaning and minimising the loss/distortion of meaning.

"Intercultural communication competence is an umbrella concept that covers intercultural awareness, intercultural sensitivity and intercultural competence" (Su, 2018, p. 218). Intercultural awareness (cognitive) is the basis for intercultural sensitivity (emotional), which in turn forms intercultural competence (behavioural). The three of them are closely related but distinct concepts (Chen, 1997). Emotional enhancement of awareness is a prerequisite for achieving the goal of competence. The focus of this study is on "intercultural sensitivity", which is related to emotion. Accordingly, intercultural sensitivity is conceptualised as "the ability to develop positive emotions of understanding and appreciation that promote appropriate and effective behaviour in intercultural communication" (Chen, 1997, p. 5). Yuen and Grosman (2009, p. 350) also state that intercultural sensitivity is "the ability to interact with other cultures and to perceive one's role in another culture". In summary, intercultural sensitivity is about "breaking down conscious and unconscious prejudices against cultural differences; it is based on understanding through recognition, inclusion through trust, and equal opportunity through appreciation" (Aydin, 2020, p. 32).

Bennett (1986) modelled intercultural sensitivity developmentally around two contexts and six orientations as seen in the Table 1.

Table 1. Developmental intercultural sensitivity model

Denial	Defence	Minimization	Acceptance	Adaptation	Integration
Ethnocentric Phases			Ethnorelativist Phases		
<i>Denial: The phase where there is little or no interest in cultural differences, where differences are not even recognised, and isolation is preferred</i>					
<i>Defence: The phase where cultural differences are recognised but react negatively to them and conformity to one's own cultural norms is expected</i>					
<i>Minimization: The phase in which, although there is a respect for cultural differences, one's own culture is seen as superior and therefore contact with the 'other' is kept to a minimum.</i>					

Acceptance: *The phase where cultural differences are recognised, validated but not yet equal*

Adaptation: *The phase where adaptation to intercultural contexts is achieved while maintaining one's own cultural codes*

Integration: *The phase where being able to interact and cooperate with other cultures seeing one's own culture as of equal value with many other cultures; to assume roles in different cultures*

"Ethnocentrism", which expresses the first context in the model, consists of phases of denial, defence and minimization. Individuals in ethnocentric phases place their own culture at the centre of their evaluation of other cultures and make judgements according to the standards of their own culture. The second context, "ethnorelativism", consists of the stages of acceptance, adaptation and integration. Individuals in ethnorelativist phases tend to interact with other cultures and experience their own culture in other cultural contexts. It is argued that as intercultural experience increases, so does sensitivity and thus intercultural competence. As can be understood, the model emphasises the creation of a broader understanding of differences and the construction of universal cultural intercultural consciousness (Aydın, 2020; Bennett, 1986; Hammer et al. 2003).

The components of intercultural sensitivity are listed as (1) self-esteem, (2) self-monitoring, (3) open-mindedness, (4) empathy, (5) interaction engagement, and (6) avoidance of judgement (Chen, 1997; Chen & Starosta, 2000). Sensitive individuals are expected to have internalised these components. To the extent that they have internalised them, they may be able to recognise and understand diversity, put themselves in the shoes of others, develop different perspectives, and have a high level of communicative competence in an effective context. This is essential for societies changing with the cultural diversity of today's world to easily overcome ethnocentrism and engaging with other cultures.

Intercultural sensitivity has the potential to positively influence attitudes towards foreign language learning as it contributes to the creation of shared meanings, interpretations and symbols and to flexibility in the face of a foreign culture. It is therefore important to consider attitudes when examining the contributions of culture to the highly complex nature of language learning. Attitudes have emerged as one of the most important emotional variables in explaining language learning outcomes because of their influence on motivation (Pavlenko, 2005). Often emphasised in psycholinguistics and educational psychology, attitude is "an individual's evaluative response to a referent or object based on beliefs or opinions" (Gardner, 1985, p. 9). This response determines preferences and shapes feelings, thoughts and behaviours. Attitudes are shaped by elements such as past experiences, social, cultural and political realities, language status, expectations, feelings of security, and so on. These elements make it easier to understand the input and also indicate that there may be changes over time. Positive attitudes can become negative and negative attitudes can become positive. According to Wenden (1991), attitudes have three components: Firstly, beliefs and perceptions about objects and situations; secondly, evaluativeness towards liking or disliking objects or situations; and thirdly, guiding learning behaviours.

Attitudes towards foreign language learning, as an individual factor influencing success and efficiency, have been addressed in research, particularly in those dimensions that relate the meaning and value that second/foreign language learners attach to language to the socio-cultural context. Papapavlou (1998, p. 16) reviewed studies on attitudes towards languages and dialects and found that the studies focused on (a) attitudes towards the foreign language and their effects on learning, (b) attitudes towards neighbouring languages and their effects on lexical borrowing, (c) attitudes towards particular social groups and the variants they speak, and (d) attitudes towards particular ethnic groups or minorities and their languages. On the other hand, models of second language acquisition take into account the nature of attitudes that may explain differences in learning outcomes. Positive attitudes facilitate access to input to the language acquisition tool, leading to high language proficiency, whereas negative attitudes hinder input, making second language acquisition difficult with high anxiety (Pavlenko, 2005). Furthermore, Tódor and Dégi (2016) found that a generally positive attitude towards languages favours interlanguage relations and integrative, dynamic multilingual thinking.

Identifying attitudes is a precursor to identifying strategies and methods to increase achievement and confidence and to meet expectations. In this way, it is possible to bridge the gap between the learner and the teacher, as well as between the learner and the content. In this respect, it has become necessary to bring a new interdisciplinary perspective and dimension to cultural studies. This perspective to be developed will contribute significantly to the formation of basic premises that will shape mental and psychological language learning processes, such as first attitude and then intercultural sensitivity in second/foreign language learning. This is why, in recent years, there has been a growing tendency to approach and analyse the subject from different angles.

Wei (2019) argues that intercultural sensitivity is mainly influenced by individual situations, learning conditions, teaching materials and intercultural communication experiences. In support of this view, studies have highlighted the relationship between the subject and various predictors. Olson and Kroeger (2001) found a relationship between intercultural sensitivity and second language acquisition and experience abroad; Straffon (2003) found a relationship between intercultural sensitivity and length of attendance at an international school; Roh (2014) found a relationship between intercultural sensitivity and level of multicultural experience; Aydoğan and Akbarov (2014), between intercultural sensitivity and attitudes towards foreign language and foreign language use; Aydoğan and Akbarov (2015), between intercultural sensitivity and foreign language proficiency, verbal ability and verbal expressiveness; Genç and Boynukara (2017), between intercultural sensitivity and social intelligence; Abaslı and Polat (2019), between intercultural sensitivity and cultural intelligence; Çiloğlu and Bardakçı (2019) found a positive relationship between intercultural sensitivity and language proficiency; Tıkız-Ertürk et al. (2023) found a positive relationship between intercultural sensitivity and intercultural competence. Engle and Engle (2004) found that study abroad develops language acquisition and intercultural sensitivity in similar ways and that there is a close relationship between the development of these

two key areas. Also, Williams (2005) found that the sensitivity of those who study abroad is significantly higher than that of those who do not and that exposure is the most important determinant of intercultural communication skills. Erdoğan and Okumuşlar (2020) found a negative relationship between intercultural sensitivity and ethnocentrism. In other words, as the level of intercultural sensitivity increases, ethnocentrism decreases. It is understood that similar results have been found in some studies that were not conducted using the relational method. For example, Penbek et al. (2012) found that intercultural sensitivity increased in parallel with the experience of participating in intercultural interactions, Laopongharn and Sercombe (2009) found that increased intercultural communication competence led to more competent language use, and Meydanlıoğlu et al. (2015) found that those who speak a foreign language and interact with different cultures have higher sensitivity; Chen and Hu (2023) found that experience abroad and grade level influence intercultural sensitivity; Çubukçu (2013) found that foreign language teacher candidates integrate language teaching objectives with culture teaching and prioritise attitudinal knowledge, tolerance and sympathy towards others.

On the other hand, there are some studies that consider the attitudes towards language learning in a relational way. Accordingly, a significant positive relationship has been found between attitudes towards language learning and academic achievement (Fakeye, 2010; İnal et al. 2005), foreign language learning strategies (Jabbari & Golkar, 2014; Saracaloğlu & Dinçer, 2017), language skill acquisition (Karatay & Kartallıoğlu, 2016), and learner autonomy (Süğümlü & Çinpolat, 2023). While a negative relationship was found between attitudes towards language learning and motivational barriers (İnalöz & Yılmaz, 2023) and success in using language skills (Karatay et al. 2018), no significant relationship was found with shyness level (Alavinia & Salmasi, 2012).

In general, the relationship between intercultural sensitivity and language learning attitude shows the limitation witnessed in the literature. However, Alisaari and Kilpi-Jakonen (2022) report that it is possible to positively influence linguistic and non-linguistic skills, including attitudes and sociolinguistic skills, by teaching intercultural competences. This suggests that a deep understanding of intercultural sensitivity can improve language learning attitudes. Ahnagari and Zamanian (2014) also reported that intercultural competence practices can increase positive attitudes towards both source and target cultures and also motivate learners to improve their language proficiency. On the other hand, the situation of Turkish origin students who came to Turkey for high school, undergraduate and graduate education for the first time within the scope of the *Big Student Project** carried out by the Ministry of National Education of the Republic of Turkey immediately after the dissolution of the Soviet Union has been the subject of many studies in various educational contexts, and the importance of providing support for language competence and cultural identity in guiding their education has been pointed out. In these studies, which are limited to linguistic frameworks such as linguistic correspondences and vocabulary, comparisons regarding the barriers in language learning processes have come to the fore. Today, both cultural and educational studies are carried out more intensively through projects and programmes produced by many institutions and organisations. As a matter of fact, linguistic and cultural belongings are the basis of developing sensitivity and attitudes towards different cultural contexts. However, the lack of such an evaluation in the related literature on Turkish origin students is seen as an important deficiency. It is thought that such a study may enable researchers, educators and policy makers to make purposeful interventions that can affect academic results and social welfare by determining a starting point. Accordingly, this study, which focuses on the relationship between intercultural sensitivity and language learning attitudes of Turkish origin students, seeks to answer the following research questions:

1. Does the intercultural sensitivity of Turkish origin students differ according to the demographic variables?
2. Does the attitudes of Turkish origin students towards foreign language learning differ according to the demographic variables?
3. Is there a significant relationship between Turkish origin students' intercultural sensitivity and their language learning attitudes?
4. Is there any effect of the Turkish origin students' intercultural sensitivity on their attitudes towards foreign language learning?

METHOD/MATERIALS

Study Model

In this study, the correlational research method was adopted to investigate the statistical relationship between intercultural sensitivity and language learning attitudes of Turkish origin students. "Correlational research is conducted either to explain the relationships between variables or to predict possible outcomes" (Fraenkel et al., 2012, p.332). For this purpose, data are collected from participants without interfering with the variables and it is accepted that the data reflect the natural relationships and degrees of relationship between the variables (Büyüköztürk et al. 2013; Fraenkel et al. 2012). In this type of research, statistical techniques such as correlation and regression are used to predict one variable from another.

* After the collapse of the Soviet Union, Turkey signed bilateral agreements with the Turkic republics and related communities to strengthen common historical and cultural ties. One of the most important projects in the field of education was the Great Student Project, which started in 1992 and continued until 2012. Thousands of students received higher education in Turkey on scholarships through the project, which aimed to raise a Turkish and pro-Turkish generation and spread Turkish language and culture.

Participants

Criterion sampling, one of the purposive sampling techniques, was used to identify the participants for the study. Criterion sampling is the inclusion of sources that have certain criteria and characteristics to meet the study objectives and can provide the richest data possible (Yıldırım & Şimşek, 2013). In this direction, the participants of the study were determined according to the criteria of having a good command of Turkish, having an undergraduate or postgraduate education in the country of residence, and being 18 years or older. Demographic information about the participants is shown in Table 2.

Table 2. Distribution of participants according to demographic variables

Variable	Group	Frequency	Percentage (%)
Gender	Female	335	65,8
	Male	174	34,2
Age	18-22 years	423	83,1
	23-26 years	42	8,3
	27 years and over	44	8,6
Nationality	Azerbaijan	115	22,6
	Kazakhstan	173	34
	Kyrgyzstan	114	22,4
	Uzbekistan	107	21
Level of Education	Undergraduate	437	85,9
	Postgraduate	72	14,1
Number of foreign languages known	Unspecified	86	16,9
	1 language	195	38,3
	2 languages	175	34,4
	3 languages or more	53	10,4
Total		509	100

65.8% of the participants were female, 34.2% were male; 83.1% were aged 18-22, 8.3% were aged 23-26 and 8.6% were aged 27 and over. In terms of nationality, they came from Kazakhstan (34%), Azerbaijan (22.6%), Kyrgyzstan (22.4%) and Uzbekistan (21%); 85.9% of them had undergraduate education and 14.1% had postgraduate education. Finally, as regards the number of languages spoken by the participants, they are divided into those who speak 1 language (38.3%), those who speak 2 languages (34.4%), those who speak 3 languages (10.4%); 16.9% of the participants did not specify the number of languages.

Data Collection

The intercultural sensitivity scale developed by Chen and Starosta (2000) and adapted into Turkish by Küllü-Sülü (2014) and the attitudes towards foreign language learning scale developed by Koç and Yastıbaş (2022) were used to collect the study data. Before using the scales, written permission was obtained from the corresponding authors. Then, both scales were transferred to the online environment and presented to the participants with a link address. Participants who had access to the link were first informed about the study and, after agreeing to participate voluntarily, responded to the relevant scales.

The *Intercultural Sensitivity Scale* was developed as a five-point likert scale with 24 items. The scale includes items such as “I enjoy communicating with people from other cultures”, “I respect the values of people from different cultures”, “I think my culture is better than other cultures”, and so on. The scale has a minimum score of 24 and a maximum score of 120. A high score indicates a high level of sensitivity. In addition, 9 items in the scale (2, 4, 7, 9, 12, 15, 18, 20, 22) are reverse scored. According to the exploratory factor analysis (EFA) results, the scale consists of five sub-dimensions. These dimensions are “interaction engagement” (7 items), “respect for cultural differences” (6 items), “interaction confidence” (5 items), “interaction enjoyment” (3 items) and “interaction attentiveness” (3 items). This scale, originally in English, was translated back and forth from English to Turkish and from Turkish to English in order to minimise losses by consulting experts and to ensure content and face validity. Then, the Cronbach's alpha reliability coefficient of the whole scale was calculated as .758 (Küllü-Sülü, 2014). In the English version of the scale, this coefficient was found to be .86, and it was found that the five-factor structure explained 37.3% of the variance (Chen & Starosta, 2000).

The *Attitudes Towards Foreign Language Learning Scale* was developed as a 34-item five-point Likert scale. The scale includes items such as “I enjoy learning a foreign language.”, “I enjoy learning a foreign language.”, “I enjoy reading sources written in a foreign language.” and so on. The scale has a minimum score of 34 and a maximum score of 170. While a low score indicates a positive attitudes, a high score indicates a negative attitudes. In addition, 13 items of the scale (7, 8, 9, 10, 11, 14, 14, 17, 24, 28, 30, 31, 32, 33) are reverse scored. The calculated Kaiser-Meyer-Olkin coefficient and the Barlett sphericity test value of the scale indicated that the scale was suitable for exploratory factor analysis (EFA). According to the EFA results, the scale consisted of five sub-dimensions. The sub-dimensions of the scale are “emotions towards foreign language learning” (10 items), “negative thoughts towards foreign language learning” (8 items), “using a foreign language for social media” (6 items), “importance given to learning

a foreign language" (5 items) and "thoughts towards foreign language lessons" (5 items). The Cronbach's alpha reliability coefficients of the sub-dimensions of the scale were .883, .867, .889, .804, .749 and .934 for the whole scale. The obtained scale explains 59.099% of the total variance (Koç & Yastıbaş, 2022).

In this study, the Cronbach's alpha reliability coefficient was found to be .79 for the Intercultural Sensitivity Scale and .92 for the Attitudes Towards Foreign Language Learning Scale. This means that the reliability of both scales is high.

Data Analysis

The SPSS 26.0 statistical package was used to analyse the data collected online as part of the study. There were no participants who gave the same answer to all items in both scales. During the analyses, normality was checked according to the skewness and kurtosis coefficients for the scales and sub-dimensions. If these values are between ± 1 , the distribution of scores is normal. For the extreme values of the scale scores, the z-standard value should be calculated and this value should be outside the range of ± 3.30 (Tabachnick & Fidell, 2013). As the skewness and kurtosis values for the sub-dimensions "interaction participation" and "interaction enjoyment" of the intercultural sensitivity scale and the sub-dimension "importance given to foreign language learning" of the attitudes towards foreign language learning scale were outside the range of ± 1 , the z values of these measures were calculated. The z-value of 12 participants was found to be outside the ± 3 range and was not included in the analysis.

Frequency and percentage distributions of demographic variables and descriptive statistics of the scales are presented in the study. The Pearson correlation method was used for the relationship between the Intercultural Sensitivity Scale and its sub-dimensions and the Attitudes Towards Foreign Language Learning Scale and its sub-dimensions. It was found that all scores were continuous and normally distributed, and the number of data ($N > 30$) was large. Turgut and Baykul (2021) state that if the correlation coefficient is greater than 0.70, there is a high correlation, if it is between 0.40-0.70, there is a medium correlation, and if it is less than 0.40, there is a weak correlation.

The effect of levels of intercultural sensitivity on attitudes towards foreign language learning was analysed using multiple linear regression analysis. Tolerance and VIF values were checked for the problem of multicollinearity, which is an important assumption for this method, and no multicollinearity problem was found. As all scale scores were normally distributed and the number of data in each group was sufficient ($N > 30$), the independent groups t-test and the one-way analysis of variance (ANOVA) method were used as parametric methods. The independent groups t-test method is used to compare continuous and normally distributed measurements of two groups while ANOVA is used to compare more than two groups (Pallant, 2007). The LSD multiple comparison method was used for significant differences as a result of ANOVA and comparisons were made at the $p < .05$ significance level for statistical analyses.

Table 3. Descriptive statistics relating to the scales scores

Scales	Scale Sub-dimensions	Lowest	Highest	Mean	sd	Skewness	Kurtosis
Intercultural Sensitivity Scale	Interaction engagement	16	35	25,87	3,09	-0,119	0,503
	Respect for cultural differences	13	30	23,61	3,29	-0,195	-0,227
	Interaction confidence	7	25	18,53	3,13	-0,247	0,249
	Interaction enjoyment	7	15	12,57	1,70	-0,42	-0,075
	Interaction attentiveness	4	15	10,51	1,73	-0,147	0,445
	Total sensitivity	65	118	91,08	8,96	-0,061	0,141
Attitudes Towards Foreign Language Learning Scale	Emotions	10	40	22,57	5,80	0,242	-0,132
	Negative thoughts	8	33	14,07	4,69	0,92	0,835
	Social media	6	28	13,50	4,11	0,221	-0,039
	Importance given	5	18	8,53	2,78	0,674	0,14
	Thoughts about lessons	5	23	10,88	3,11	0,137	-0,092
	Total attitudes	34	120	69,56	16,34	0,177	-0,342

The total and sub-dimensions scores for the scales used were obtained by summing the items. The total score of the Intercultural Sensitivity Scale ranged from 65 to 118 and the mean score was 91.08. The mean scores for the sub-dimensions of the same scale were obtained as "interaction engagement" ($X=25.37$), "respect for cultural differences" ($X=23.61$), "interaction confidence" ($X=18.53$), "interaction enjoyment" ($X=12.57$) and "interaction attentiveness" ($X=10.51$). The total score of the Attitudes Towards Foreign Language Learning Scale ranged from 34 to 120 and the mean score was 69.56. The means for the sub-dimensions of the same scale were obtained as "emotions towards foreign language learning" ($X=22.57$), "negative thoughts towards foreign language learning" ($X=14.07$), "using a foreign language for social media" ($X=13.50$), "thoughts towards foreign language lessons" ($X=10.88$) and "importance given to learning a foreign language" ($X=8.53$). As the skewness and kurtosis values for both scales and sub-dimensions are between ± 1 , the distribution of the scores is normal.

FINDINGS

Firstly, the relationship between intercultural sensitivity and attitudes towards foreign language learning of Turkish origin students and demographic variables was examined. The findings regarding the comparison of Intercultural Sensitivity Scale scores according to demographic variables are given in Table 4.

Table 4. Comparison of intercultural sensitivity scale scores according to demographic variables

Variable	Group	N	Int. eng.	Respect cult. dif.	Int. conf.	Int. enjoy.	Int. attent.	Total sens.
			X±sd	X±sd	X±sd	X±sd	X±sd	X±sd
Gender	Female	335	25,9±3,1	23,9±3,1	18,4±3,1	12,6±1,7	10,4±1,7	91,3±8,6
	Male	174	25,8±3,2	23,1±3,5	18,7±3,1	12,5±1,7	10,7±1,8	90,7±9,7
	t		0,537	2,708	-1,136	0,813	-1,373	0,664
	p		0,592	0,007*	0,256	0,416	0,17	0,507
Age	18-22 years old (a)	149	25,6±3,1	23,3±3,2	18,3±3,1	12,5±1,7	10,4±1,7	90,2±8,8
	23-26 years old (b)		26,4±2,6	24±3,6	19,1±2,6	12,8±1,5	10,7±1,9	93±8,3
	27 years and over (c)	151	27,6±2,6	25,7±2,9	20,2±3	13,3±1,3	11,2±2,1	97,9±7,7
	Test statistic		8,486	10,805	7,893	5,274	4,887	16,973
	p		,000*	,000*	,000*	0,005*	0,008*	,000*
	Difference**		c>a	c>a,b	c>a	c>a	c>a	c>b>a
Nationality	Azerbaijan (a)	115	26,9±2,7	25,3±2,8	19±3,4	12,9±1,6	11,1±1,9	95,2±8,5
	Kazakhstan (b)	173	25,4±3	22,6±3,5	18,7±3	12,4±1,8	10,3±1,6	89,4±8,8
	Kyrgyzstan (c)	114	25,7±3,2	23,6±2,6	17,7±3,4	12,4±1,6	10,2±1,6	89,7±8,3
	Uzbekistan (d)	107	25,7±3,4	23,3±3,4	18,6±2,7	12,7±1,7	10,6±1,7	91±9,1
	F		5,598	17,023	3,491	3,207	6,708	11,853
	p		0,001*	,000*	0,016*	0,023*	,000*	,000*
	Difference**		a>b,c,d	a>b,c,d; c>b	a,b,d>c	a>b,c	a>b,c,d	a>b,c,d
Level of Education	Undergraduate	437	25,8±3,1	23,6±3,3	18,5±3,1	12,6±1,7	10,4±1,7	90,9±9
	Graduate	72	26,2±3,3	23,6±3,3	18,8±3,2	12,4±1,5	11,1±1,8	92,2±8,9
	t		-0,912	-0,089	-0,77	0,664	-3,21	-1,107
	p		0,362	0,929	0,442	0,507	0,001*	0,269
Number of Foreign Languages Known	Unspecified (a)	79	25,5±3	23,6±3,5	17,9±3,3	12,6±1,8	10,6±1,8	90,2±9,7
	1 language (b)	97	25,8±3,1	23,8±3,1	18,3±3,1	12,7±1,6	10,6±1,7	91,2±8,6
	2 languages (c)	76	25,9±3,1	23,4±3,4	18,7±3,1	12,4±1,7	10,3±1,7	90,7±9
	3 languages and over (d)	48	26,6±3,2	23,5±3,4	19,8±2,8	12,8±1,9	10,9±1,6	93,6±8,6
	F		1,393	0,393	4,88	1,077	1,974	1,856
	p		0,244	0,758	0,002*	0,358	0,117	0,136
	Difference**				d>a,b,c; c>a			

*p<.05; **LSD multiple comparison; X=Average; sd=standard deviation

According to the gender of the participants, a significant difference was obtained only between the scores of the “respect for cultural differences” sub-dimension of intercultural sensitivity, and the “respect for cultural differences” sub-dimension average of the women was higher than that of the men (p<.05).

A statistically significant difference was found between the total intercultural sensitivity scores and all sub-dimension scores according to the age of the participants (p>.05). Accordingly, the mean scores of the participants aged 27 years and above for the sub-dimensions of “interaction engagement”, “interaction confidence”, “interaction enjoyment” and “interaction attentiveness” were higher than those of the participants aged 18-22 years. In addition, the mean scores of participants aged 27 and over on “respect for cultural differences” are higher than those of participants aged 18-22 and 23-26. Finally, the total mean score for intercultural sensitivity of the participants aged 27 years and over was higher than that of the participants aged 23-26 years and 18-22 years, and the total mean score of the participants aged 23-26 years was higher than that of the participants aged 18-22 years.

There was a statistically significant difference between the total intercultural sensitivity scores and all sub-dimension scores according to the nationality of the participants (p>.05). For all ANOVA results, the source of the difference was compared with LSD post hoc method. The mean of “interaction engagement”, “respect for cultural differences”, “interaction attentiveness” sub-dimensions and total sensitivity score of the participants from Azerbaijan is higher than all participants from other nationalities. In addition, the mean score of “respect for cultural differences” of Kyrgyzstan national participants was higher than Kazakhstan national participants. The mean score of “interaction confidence” of Kyrgyzstan national participants is lower than all participants of other nationalities. The mean score of “interaction enjoyment” of the participants from Azerbaijan was higher than the participants from Kazakhstan and Kyrgyzstan.

According to the education level of the participants, there is a significant difference only between the “interaction attentiveness” sub-dimension scores (p<.05) and the mean score of the “interaction attentiveness” sub-dimension of the participants at the graduate level was higher than the participants at the undergraduate level.

According to the number of foreign languages learnt, there is a significant difference only between the “interaction confidence” sub-dimension scores (p<.05). The mean score of the participants who know 3 or more languages is higher than the participants

who know 2 languages, 1 language and the participants who do not specify; the mean score of the participants who know 2 languages is higher than the participants who do not specify.

The findings regarding the comparison of Attitudes Towards Foreign Language Learning Scale scores according to demographic variables are given in Table 5.

Table 5. Comparison of attitudes scores towards foreign language learning scale according to demographic variables

Variable	Group	N	Emotions	Negative thoughts	Social media	Attention given	Thoughts about lessons	Total attitudes
			X $\pm$ sd	X $\pm$ sd	X $\pm$ sd	X $\pm$ sd	X $\pm$ sd	X $\pm$ sd
Gender	Female	161	22,3 $\pm$ 5,6	13,7 $\pm$ 4,4	13,3 $\pm$ 4,1	8,3 $\pm$ 2,6	10,6 $\pm$ 3,1	68,3 $\pm$ 15,4
	Male	139	23,1 $\pm$ 6,2	14,8 $\pm$ 5,1	13,8 $\pm$ 4,2	9 $\pm$ 3	11,4 $\pm$ 3,2	72 $\pm$ 17,8
	t		-1,52	-2,476	-1,209	-2,47	-2,674	-2,486
	p		0,129	0,014*	0,227	0,014*	0,008*	0,013*
Age	18-22 years old (a)	149	23 $\pm$ 5,6	14,4 $\pm$ 4,7	13,7 $\pm$ 4	8,7 $\pm$ 2,8	11,2 $\pm$ 3	71 $\pm$ 15,9
	23-26 years old (b)		20,8 $\pm$ 6,1	13,6 $\pm$ 4,6	12,7 $\pm$ 3,8	8 $\pm$ 3,3	10,1 $\pm$ 3,2	65,3 $\pm$ 17,3
	27 years and over (c)	151	20 $\pm$ 6,8	11,8 $\pm$ 4,1	12,2 $\pm$ 4,7	7,3 $\pm$ 2,1	8,8 $\pm$ 2,9	60 $\pm$ 16,3
	Test statistic		7,646	6,576	3,546	6,193	13,833	10,949
	p		0,001*	0,002*	0,03*	0,002*	0*	0*
	Difference**		a>b,c	a>c	a>c	a>c	a>b>c	a>b,c
Nationality	Azerbaijan (a)	115	22,8 $\pm$ 6,6	12,1 $\pm$ 3,9	13,3 $\pm$ 4,4	7,9 $\pm$ 2,6	10,3 $\pm$ 3,5	66,2 $\pm$ 17,5
	Kazakhstan (b)	173	22,8 $\pm$ 5,5	14,7 $\pm$ 4,7	13,8 $\pm$ 3,8	8,8 $\pm$ 2,8	11,7 $\pm$ 2,9	71,9 $\pm$ 15,8
	Kyrgyzstan (c)	114	23,4 $\pm$ 5,5	14,8 $\pm$ 4,4	14 $\pm$ 4,3	9 $\pm$ 2,6	10,8 $\pm$ 2,8	72 $\pm$ 15,5
	Uzbekistan (d)	107	21,1 $\pm$ 5,4	14,4 $\pm$ 5,2	12,7 $\pm$ 4	8,3 $\pm$ 2,9	10,3 $\pm$ 3	66,8 $\pm$ 16
	F		3,419	9,769	2,428	4,532	7,615	4,751
	p		0,017*	0*	0,065	0,004*	0*	0,003*
	Difference**		a,b,c>d	a<b,c,d		b,c>a; c>d	b>a,c,d	b,c>a,d
Level of Education	Graduate	437	22,7 $\pm$ 5,7	14,1 $\pm$ 4,7	13,5 $\pm$ 4,2	8,6 $\pm$ 2,8	11 $\pm$ 3,1	69,8 $\pm$ 16,3
	Undergraduate	72	22 $\pm$ 6,1	14,2 $\pm$ 4,9	13,6 $\pm$ 3,9	8,4 $\pm$ 2,8	10,2 $\pm$ 3	68,3 $\pm$ 16,7
	t		0,884	-0,214	-0,157	0,525	2,113	0,702
	p		0,377	0,83	0,876	0,6	0,035*	0,483
Number of foreign languages known	Unspecified (a)	79	24,3 $\pm$ 6,1	14,7 $\pm$ 4,9	15 $\pm$ 4,5	9 $\pm$ 3	11,4 $\pm$ 3,5	74,3 $\pm$ 17,6
	1 language (b)	97	22,8 $\pm$ 5,8	14,1 $\pm$ 4,7	13,4 $\pm$ 4,2	8,4 $\pm$ 2,9	10,7 $\pm$ 3,1	69,3 $\pm$ 16,6
	2 languages (c)	76	22,4 $\pm$ 5,4	14,1 $\pm$ 4,6	13,2 $\pm$ 3,7	8,6 $\pm$ 2,6	10,9 $\pm$ 2,8	69,1 $\pm$ 14,9
	3 languages and over (d)	48	19,7 $\pm$ 5,5	13,1 $\pm$ 4,6	12,7 $\pm$ 4	8,1 $\pm$ 2,5	10,7 $\pm$ 3,5	64,2 $\pm$ 16,3
	F		7,341	1,275	4,788	1,421	1,012	4,453
	p		0*	0,282	0,003*	0,236	0,387	0,004*
	Difference**		a>b,c>d		a>b,c,d			a>b,c>d

* $p<.05$; **LSD multiple comparison; X=Average; sd=standard deviation

A significant difference was found between the participants' overall attitudes scores towards foreign language learning and the sub-dimension scores of "negative thoughts towards foreign language learning" and "thoughts towards foreign language lessons" according to their gender ($p<.05$) is higher. This finding means that women's attitudes towards foreign language learning are more positive in terms of overall attitudes scores and related sub-dimensions.

A statistically significant difference was found between the total scores of the participants' attitudes towards foreign language learning and all subscale scores according to their ages ($p>.05$). The total scores of the participants aged 18-22 and above in terms of attitudes towards foreign language learning and the mean scores of the "emotions towards foreign language learning" subscale were higher than the participants aged 23 and above. In this case, it is understood that the general attitudes of the participants aged 23 and over towards foreign language learning and their attitudes towards the sub-dimension of "emotions towards foreign language learning" are more positive. The mean scores of the participants between the ages of 18-22 on the sub-dimensions of "negative thoughts towards foreign language learning", "using foreign language for social media" and "importance given to learning a foreign language" are higher than the participants aged 27 and above. In other words, the attitudes of the participants aged 27 and above in the related sub-dimensions are more positive than the participants aged 18-22. Finally, the mean scores of the participants between the ages of 18-22 on the sub-dimension "thoughts about foreign language lessons" were higher than those of the participants aged 23 and above, and the mean scores of the participants between the ages of 23-26 were higher than those of the participants aged 27 and above. This finding shows that the attitudes in the relevant sub-dimension become more positive as the age of the participants increases.

There was a statistically significant difference between the total scores of participants' attitudes towards foreign language learning and all sub-dimension scores except "using foreign language for social media" according to their nationality ($p>.05$). For all ANOVA results, the source of the difference was compared using the LSD post hoc method. The mean score of the participants

from Uzbekistan in the subscale “emotions towards foreign language learning” was lower than all participants from other nationalities. This finding means that the attitudes of the participants from Uzbekistan are more positive than the others in this sub-dimension. The mean score of the participants from Azerbaijan on the sub-dimension “negative thoughts about learning a foreign language” is lower than that of the participants from other nationalities. According to this finding, the attitudes of the participants from Kazakhstan, Kyrgyzstan and Uzbekistan are more positive than those of the participants from Azerbaijan. The mean scores of the participants from Kazakhstan and Kyrgyzstan on the sub-dimension “importance given to foreign language learning” are higher than those of the participants from Azerbaijan, and the mean scores of the participants from Kyrgyzstan are higher than those of the participants from Uzbekistan. In other words, the attitudes of the participants from Azerbaijan are more positive than the attitudes of the participants from Kazakhstan and Kyrgyzstan; the attitudes of the participants from Uzbekistan are more positive than the attitudes of the participants from Kyrgyzstan. The mean score of the participants from Kazakhstan in the sub-dimension “thoughts about foreign language lessons” is higher than that of the participants from other nationalities. Similarly, the attitudes of the participants from Azerbaijan, Kyrgyzstan and Uzbekistan are more positive in this dimension than those of the participants from Kazakhstan. Finally, the total mean scores of the participants from Kazakhstan and Kyrgyzstan were higher than the participants from Azerbaijan and Uzbekistan. Therefore, the attitudes of the participants from Azerbaijan and Uzbekistan are more positive than those of the participants from Kazakhstan and Kyrgyzstan.

According to the level of education of the participants, there is only a significant difference ($p < .05$) between the scores of the sub-dimension “attitudes towards foreign language lessons”, and the mean score of the participants at the undergraduate level is higher than that of the participants at the graduate level. This finding shows that the attitudes of the participants at the graduate level are more positive in the relevant sub-dimension.

According to the number of foreign languages known, there is a significant difference between the total scores of attitudes towards foreign language learning and the sub-dimension scores of “emotions towards foreign language learning” and “using foreign language for social media” ($p < .05$). The mean scores of the participants who did not indicate the number of foreign languages were higher than those who indicated that they knew 1 or more languages, and the mean scores of the participants who knew 1 and 2 languages were higher than those who knew 3 languages. This finding shows that the attitudes of the participants who knew 3 languages were more positive in terms of general attitudes and the sub-dimension “emotions towards foreign language learning”. Finally, the mean score of the participants who did not indicate the number of foreign languages in the sub-dimension “using foreign language for social media” was higher than the mean score of the participants who knew 1 or more foreign languages. In other words, the participants who indicated that they have at least one or more foreign languages have more positive attitudes than the participants who did not indicate any foreign languages.

Following the demographic evaluations, it was analysed whether there was a statistically significant relationship between intercultural sensitivity and attitudes towards foreign language learning among Turkish origin students. In this context, the relationship between the total scores and the sub-dimensions of the Intercultural Sensitivity Scale and the Attitudes Towards Foreign Language Learning Scale was analysed using the Pearson correlation method (see Table 6).

Table 6. Pearson correlation table between scale scores

	Emotions	Negative thoughts	Social media	Attention given	Thoughts about lessons	Total attitudes
Interaction engagement	-,401**	-,422**	-,365**	-,511**	-,354**	-,509**
Respect for cultural differences	-,376**	-,554**	-,291**	-,459**	-,415**	-,523**
Interaction confidence	-,424**	-,299**	-,327**	-,323**	-,309**	-,432**
Interaction enjoyment	-,403**	-,460**	-,300**	-,411**	-,344**	-,486**
Interaction attentiveness	-,170**	-,170**	-,136**	-,314**	-,129**	-,221**
Total sensitivity	-,534**	-,574**	-,430**	-,596**	-,473**	-,654**

** $p < .01$; * $p < .05$

A statistically significant and negative relationship was found between the total score and sub-dimensions of the Intercultural Sensitivity Scale and the total score and all sub-dimensions of the Attitudes Towards Foreign Language Learning Scale ($p < .01$). In other words, as intercultural sensitivity total scores increase, attitudes towards foreign language learning scores decrease or as sensitivity total scores decrease, attitudes towards foreign language learning scores increase. However, while evaluating the negative relationship here, it should be noted that a low score on the attitudes scale indicates a positive attitude and a high score indicates a negative attitude. Therefore, students' high intercultural sensitivity means that their attitudes towards foreign language learning are positive. The relationship between total sensitivity scores and total attitudes towards foreign language learning scores was -0.654 , and the relationship between the attitude scale sub-dimensions was between -0.430 and -0.596 . The relationship between the sub-dimension of “interaction engagement” and total scores of attitudes towards language learning was -0.509 and the relationship between the sub-dimensions of the attitudes scale was between -0.354 and -0.511 . The relationship between the sub-dimension of “respect for cultural differences” and total scores of attitudes towards foreign language learning was -0.523 and the relationship between the sub-dimensions of the attitudes scale was between -0.291 and -0.554 ; the relationship between the sub-dimension of “interaction confidence” and total scores of attitudes towards foreign language learning was -0.432 and the relationship between the sub-dimensions of the attitudes scale was between -0.291 and -0.554 ; the relationship between the sub-dimension of “interaction confidence” and total scores of attitudes towards foreign language

learning was -0.432 and the relationship between the sub-dimensions of the attitudes scale was -0.432. 432 and the relationship between the attitudes scale sub-dimensions was between -0.299 and -0.424; the relationship between the “interaction enjoyment” sub-dimension and the total scores of attitudes towards foreign language learning was -0.486 and the relationship between the attitudes scale sub-dimensions was between -0.300 and -0.460. Finally, the relationship between the “interaction attentiveness” sub-dimension and the total scores of attitudes towards foreign language learning was -0.221 and the relationship between the attitudes scale sub-dimensions was between -0.129 and -0.314.

In this study, the effect of Turkish origin students' intercultural sensitivity on their attitudes towards foreign language learning was analysed using the multiple linear regression analysis method. This method analyses the effect of more than one independent variable or predictor on the dependent variable. For regression analysis, the dependent variable should be continuous and normally distributed, and there should be no problem of multicollinearity between the independent variables. Accordingly, all tolerance values are expected to be greater than 0.10 and VIF values are expected to be less than 10 (Pallant, 2007). Since the tolerance values are between 0.663-0.883 and the VIF values are between 1.132-1.539, there is no multicollinearity problem.

Table 7. Regression table for the effect of intercultural sensitivity on attitudes towards foreign language learning

Dependent Variable	Independent Variables	B	S.H.	β	t	p	ANOVA	R ²
Total attitudes	Fixed	178,178	5,727		31,111	,000	F(5,503)=80.043; p<.001	0.443
	Interaction engagement	-1,198	0,216	-0,227	-5,548	,000		
	Respect for cultural differences	-1,417	0,194	-0,285	-7,314	,000		
	Interaction confidence	-0,901	0,2	-0,172	-4,494	,000		
	Interaction enjoyment	-1,91	0,383	-0,199	-4,993	,000		
	Interaction attentiveness	-0,329	0,334	-0,035	-0,987	0,324		
Emotions towards foreign language learning	Fixed	53,269	2,279		23,376	,000	F(5,503)=42.915; p<.001	0.299
	Interaction engagement	-0,307	0,086	-0,164	-3,568	,000		
	Respect for cultural differences	-0,292	0,077	-0,165	-3,784	,000		
	Interaction confidence	-0,446	0,08	-0,241	-5,593	,000		
	Interaction enjoyment	-0,562	0,152	-0,165	-3,691	,000		
	Interaction attentiveness	-0,052	0,133	-0,016	-0,392	0,695		
Negative thoughts towards foreign language learning	Fixed	41,955	1,721		24,38	,000	F(5,503)=64.232; p<.001	0.390
	Interaction engagement	-0,222	0,065	-0,147	-3,425	0,001		
	Respect for cultural differences	-0,544	0,058	-0,381	-9,35	,000		
	Interaction confidence	-0,057	0,06	-0,038	-0,943	0,346		
	Interaction enjoyment	-0,614	0,115	-0,223	-5,345	,000		
	Interaction attentiveness	-0,048	0,1	-0,018	-0,483	0,629		
Using foreign language for social media	Fixed	31,146	1,733		17,974	,000	F(5,503)=24.261; p<.001	0.194
	Interaction engagement	-0,279	0,065	-0,21	-4,276	,000		
	Respect for cultural differences	-0,137	0,059	-0,109	-2,332	0,02		
	Interaction confidence	-0,225	0,061	-0,171	-3,708	,000		
	Interaction enjoyment	-0,241	0,116	-0,100	-2,081	0,038		
	Interaction attentiveness	0	0,101	0	0,004	0,997		
Importance given to foreign language learning	Fixed	26,241	1,029		25,505	,000	F(5,503)=60.851; p<.001	0.377
	Interaction engagement	-0,249	0,039	-0,277	-6,413	,000		
	Respect for cultural differences	-0,191	0,035	-0,226	-5,486	,000		
	Interaction confidence	-0,044	0,036	-0,05	-1,226	0,221		
	Interaction enjoyment	-0,271	0,069	-0,166	-3,94	,000		
	Interaction attentiveness	-0,242	0,06	-0,151	-4,031	,000		
Thoughts about foreign language lessons	Fixed	25,567	1,272		20,096	,000	F(5,503)=32.068; p<.001	0.242
	Interaction engagement	-0,141	0,048	-0,14	-2,942	0,003		
	Respect for cultural differences	-0,253	0,043	-0,268	-5,888	,000		
	Interaction confidence	-0,129	0,045	-0,13	-2,893	0,004		
	Interaction enjoyment	-0,222	0,085	-0,121	-2,614	0,009		
	Interaction attentiveness	0,012	0,074	0,007	0,166	0,868		

All multiple linear regression models regarding the effect of the sub-dimensions of the Intercultural Sensitivity Scale on the overall score of the Attitudes Towards Foreign Language Learning Scale ($F_{(5,503)}=80.043$, $p<.001$), the scores in the sub-dimensions of “emotions towards foreign language learning” ($F_{(5,503)}=42.915$, $p<.001$), “negative thoughts towards foreign language learning” ($F_{(5,503)}=64.232$, $p<.001$), “using a foreign language for social media” ($F_{(5,503)}=24.261$, $p<.001$), “importance given to learning a foreign language” ($F_{(5,503)}=60.851$, $p<.001$) and “thoughts towards foreign language lessons” ($F_{(5,503)}=32.068$, $p<.001$) were statistically significant. The R^2 value shows the variance explained in the dependent variable (Pallant, 2007). According to this, 44.3% of the total scores of attitudes towards foreign language learning, 29.9% of the scores of “emotions towards foreign language learning”, 39% of the scores of “negative thoughts towards foreign language learning”, 19.4% of the scores of “using a foreign language for social media”, 37.7% of the scores of “importance given to foreign language learning” and 24.2% of the scores of “thoughts towards foreign language lessons” are explained by the sub-dimensions of the Intercultural Sensitivity Scale.

The sub-dimensions of the Intercultural Sensitivity Scale “interaction engagement” ($B=-1.198$, $t=-5.548$, $p<.05$), “respect for cultural differences” ($B=-1.417$, $t=-7.314$, $p<.05$), “interaction confidence” ($B=-0.901$, $t=-4.494$, $p<.001$) and “interaction

enjoyment" ($B=-1.91$, $t=-4.493$, $p<.001$) were found to be significant predictors of general attitudes towards foreign language learning scores. The effect of these sub-dimension scores obtained as significant predictors on the participants' general attitudes towards foreign language learning is negative. The β coefficient indicates the order of importance of the independent variables in the model. Among these significant predictors, the sub-dimension with the highest effect was "respect for cultural differences", while the one with the lowest effect was "interaction confidence". The sub-dimension "interaction attentiveness" was not a significant predictor of general attitudes towards foreign language learning ($p>.05$).

The sub-dimensions of the Intercultural Sensitivity Scale "interaction engagement" ($B=-0.307$, $t=-3.568$, $p<.05$), "respect for cultural differences" ($B=-0.292$, $t=-3.784$, $p<.05$), "interaction confidence" ($B=-0.446$, $t=-5.593$, $p<.001$) and "interaction enjoyment" ($B=-0.562$, $t=-3.961$, $p<.001$) were found to be significant predictors of the "emotions towards foreign language learning" sub-dimension scores. The effect of these sub-dimension scores obtained as significant predictors on participants' "emotions towards foreign language learning" is negative. Among these significant predictors, the sub-dimension with the highest effect was "interaction confidence" and the one with the lowest effect was "interaction engagement". The sub-dimension "interaction attentiveness" was not a significant predictor for the sub-dimension "emotions towards foreign language learning" ($p>.05$).

The sub-dimensions of the Intercultural Sensitivity Scale "interaction engagement" ($B=-0.222$, $t=-3.425$, $p<.05$), "respect for cultural differences" ($B=-0.544$, $t=-9.35$, $p<.05$) and "interaction enjoyment" ($B=-0.614$, $t=-5.345$, $p<.001$) were found to be significant predictors of the sub-dimension scores of "negative thoughts towards foreign language learning". The effect of these sub-dimension scores obtained as significant predictors of participants' "negative thoughts towards foreign language learning" is negative. Among these significant predictors, the sub-dimension with the highest effect was "respect for cultural differences", while the sub-dimension with the lowest effect was "interaction engagement". The sub-dimensions "interaction confidence" and "interaction attentiveness" were not significant predictors of the sub-dimension "negative thoughts towards foreign language learning" ($p>.05$).

The sub-dimensions of the Intercultural Sensitivity Scale "interaction engagement" ($B=-0.279$, $t=-4.276$, $p<.05$), "respect for cultural differences" ($B=-0.137$, $t=-2.332$, $p<.05$), "interaction confidence" ($B=-0.225$, $t=-3.708$, $p<.001$) and "interaction enjoyment" ($B=-0.241$, $t=-2.081$, $p<.05$) were found to be significant predictors of the sub-dimension scores of "using foreign language for social media". The effect of these sub-dimension scores obtained as significant predictors of "using foreign language for social media" is negative. Among these significant predictors, the sub-dimension with the highest effect was "interaction engagement", while the least effective was "interaction enjoyment". The sub-dimension "interaction attentiveness" was not a significant predictor of "using foreign language for social media" ($p>.05$).

The sub-dimensions of the Intercultural Sensitivity Scale "interaction engagement" ($B=-0.249$, $t=-6.413$, $p<.05$), "respect for cultural differences" ($B=-0.191$, $t=-5.486$, $p<.05$), "interaction enjoyment" ($B=-0.271$, $t=-3.94$, $p<.001$) and "interaction attentiveness" ($B=-0.242$, $t=-4.031$, $p<.05$) were found to be significant predictors of the "importance given to foreign language learning" sub-dimension scores. The effect of these sub-dimension scores obtained as significant predictors on the "importance given to foreign language learning" by the participants is negative. Among these significant predictors, the sub-dimension with the highest effect was "interaction engagement", while the least effective was "interaction attentiveness". The sub-dimension "interaction confidence" was not a significant predictor for the sub-dimension "importance given to foreign language learning" ($p>.05$).

The sub-dimensions of the Intercultural Sensitivity Scale "interaction engagement" ($B=-0.141$, $t=-2.942$, $p<.05$), "respect for cultural differences" ($B=-0.253$, $t=-5.888$, $p<.05$), "interaction confidence" ($B=-0.129$, $t=-2.893$, $p<.01$) and "interaction enjoyment" ($B=-0.222$, $t=-2.614$, $p<.05$) were found to be significant predictors of the "thoughts about foreign language lessons" sub-dimension scores. The effect of these sub-dimension scores obtained as significant predictors on participants' "thoughts towards foreign language lessons" is negative. Among these significant predictors, the sub-dimension with the highest effect was "respect for cultural differences", while the least effective was "interaction enjoyment". The sub-dimension "interaction attentiveness" was not a significant predictor for the sub-dimension "thoughts about foreign language lessons" ($p>.05$).

DISCUSSION, CONCLUSION AND SUGGESTIONS

In this study focusing on Turkish origin students, their intercultural sensitivity and attitudes towards foreign language learning were firstly analysed in terms of demographic variables. It was found that gender had no effect on the overall intercultural sensitivity scores and that there was a difference in favour of women only in the "respect for cultural differences" sub-dimension. This result is in line with studies showing that gender is not a determining variable for intercultural sensitivity (Abaslı & Polat, 2019; Çiloğlu & Bardakçı, 2019; Erdoğan & Okumuşlar, 2020; Meydanlıoğlu et al. 2015; Hammer et al. 2003; Tıkız-Ertürk et al. 2023; Yılmaz & Göçen, 2013). There are also different gender-related findings in the literature. For example, in some studies, women (Holm et al. 2009; Mellizo, 2017; Petrovic & Zlatkovic, 2009) and in some other studies, men (Ghamarnia et al. 2016) were found to have higher total intercultural sensitivity scores. Furthermore, Meydanlıoğlu et al. (2015) found a significant difference in the "interaction confidence" and "interaction enjoyment" sub-dimensions, Chen and Hu (2023) found a significant difference in favour of men in the "interaction confidence" sub-dimension, and Boudouaia et al. (2022) found a significant difference in favour of women in the "interaction enjoyment" sub-dimension. In addition, Erdoğan and Okumuşlar (2020) found a significant difference in favour of men on ethnocentrism. Boudouaia et al. (2022) state that considering the role of gender can provide a basis

for analysing the strengths and weaknesses of men and women in teaching. When the gender variable was examined within the framework of attitudes towards foreign language learning, it was determined that women had more positive attitudes in the total attitudes scores and the sub-dimensions of “negative thoughts towards foreign language learning” and “thoughts towards foreign language lessons”. This result is similar to the results of some studies (İnal et al. 2005; Karahan, 2007), while it differs from the studies that reveal that gender is not effective in attitude (Alavinia & Salmasi, 2012; Fakeye, 2010; Saracaloğlu & Dinçer, 2017; Yaprak & Uslu-Üstten, 2021). Edlund et al. (2007, pp. 56-60) state that women and men have different tendencies in terms of the way they use linguistic knowledge, communicate and utilise language styles. Ellis (1994) explains that women have a more positive attitudes towards second language learning because they are more open to different language forms.

According to the age variable, a significant difference was found in the overall intercultural sensitivity scores and all sub-dimension scores in favour of participants aged 27 and over. In addition, it is a very important finding that the overall level of sensitivity increases in parallel with age. Erdoğan and Okumuşlar (2020) also found that age created a significant difference in intercultural sensitivity levels. However, there are also studies in the literature that revealed that age has no effect on intercultural sensitivity (Abaslı & Polat, 2019; Genç & Boynukara 2017; Ghamarnia et al. 2016; Hammer et al. 2003; Tıkız-Ertürk et al. 2023). When the effect of the age variable on attitudes towards foreign language learning was assessed, a significant difference was found between the overall attitudes scores and all the sub-dimension scores. Participants aged 23 and over had more positive attitudes towards foreign language learning in general and in the sub-dimension of “emotions towards foreign language learning”; participants aged 27 and over had more positive attitudes towards the sub-dimensions of “negative thoughts towards foreign language learning”, “using a foreign language for social media” and “importance given to foreign language learning”. For the sub-dimension “thoughts about foreign language lessons”, attitudes became more positive with increasing age. These findings suggest that attitudes towards language learning develop largely in parallel with age. In support of this, Saville-Troike (2012) states that although it is accepted that children are more successful language learners due to their ability to create and act in acquisition contexts, adults tend to understand second/foreign languages better because they have more analytical skills, learning capacity, pragmatic skills, mother tongue knowledge and real world knowledge. However, Karahan (2007) states that the attitudes of those who learn the language at an earlier age are more positive than those who learn it at a later age, while Yaprak and Uslu-Üstten (2021) found that age has no effect on attitude.

According to the nationality variable, a significant difference in favour of Azerbaijan was found in the overall intercultural sensitivity scores and in all sub-dimension scores. It is suggested that the main reason for the higher scores of the Azerbaijani participants may be related to the fact that the data collection instruments were in Turkish. Although the participants from Kazakhstan, Kyrgyzstan and Uzbekistan know Turkish, they are not as dominant in Turkish as the participants from Azerbaijan. Contrary to this study, nationality was found to have no significant effect on the level of intercultural sensitivity (Abaslı & Polat, 2019; Erdoğan & Okumuşlar, 2020). Ghamarnia et al. (2016) also found no significant relationship between ethnicity and intercultural sensitivity. On the other hand, Erdoğan and Okumuşlar (2020) found that the intercultural sensitivity of students from Europe and students from Asia and Africa were similar while the level of Turkish students was lower. When analysing the effect of the nationality variable on attitudes towards foreign language learning, a significant difference was found in the overall attitudes scores and in all sub-dimensions except “using foreign language for social media”. In the sub-dimensions with significant differences, the countries with more positive attitudes differ. Nationality is a concept that reflects linguistic and cultural belonging. Therefore, closeness or distance to the target language can shape attitudes and affect performance. What they understand by foreign language learning, which language(s) they consider to be foreign and their self-reports may explain the lack of significance between the overall attitudes scores. For example, if the participants answered the scales with Russian in mind, the Russian influence in the Turkish communities may lead to more common attitudes in the countries covered, resulting in no difference in the overall attitudes scores.

According to the variable level of education, only the scores of the sub-dimension “interaction attentiveness” were significantly higher in favour of the participants with postgraduate education. Considering that the level of education develops perceptions, opinions and perspectives, it is expected to be a determinant of intercultural sensitivity. However, contrary to expectations, the results of this study are consistent with studies that show that the level of education is not a determinant for intercultural sensitivity (Abaslı & Polat, 2019; Bayles, 2009; Boudouaia et al. 2022). Nevertheless, it can be said that the “interaction attentiveness” sub-dimension, which focuses on obtaining more information and more details, contributed to the confirmation of the expectation, albeit slightly. When the educational level variable was examined in relation to attitudes towards foreign language learning, it was found that it did not affect the overall attitudes scores; there was only a difference in favour of graduate students in the sub-dimension “thoughts about foreign language lessons”. It is thought that attitudes towards foreign language learning are directly related to educational level. While it is expected that increased awareness might have an effect on attitudes towards learning a new language and attitudes towards learning that language, it is noteworthy that the research result does not confirm this.

According to the variable of the number of foreign languages known, it was found that only among the scores of the sub-dimension “interaction confidence”, those who know three languages and more have higher scores than the others, and those who know two languages have higher scores than those who do not specify a language. Knowing a foreign language is an important part of building social skills and self-confidence, and as it also means experiencing the culture to which that language belongs, it is expected to increase cultural sensitivity. The developmental nature of intercultural sensitivity also increases efficiency and self-

confidence in parallel with the frequency of communication (Bennett, 1986). The difference in “interaction confidence” is therefore normal. However, it is worth noting that the number of known foreign languages has no effect on the total score for sensitivity. This result is in line with Tıkız-Ertürk et al. (2023). On the other hand, Meydanlıoğlu et al. (2015) found that the total score of intercultural sensitivity and the subdimension scores of “respect for cultural differences”, “interaction confidence” and “interaction enjoyment” were higher in the case of foreign language proficiency. Chen and Hu (2023) found that the mean sensitivity scores of students who passed the foreign language proficiency exam were higher than those of those who did not pass; a significant difference was found in the “interaction confidence” dimension. Olson and Kroeger (2001) also found that foreign language speakers have higher intercultural sensitivity. When the variable of the number of foreign languages learnt was evaluated according to the attitudes towards foreign language learning, it affected the total attitudes scores and the sub-dimension scores of “emotions towards foreign language learning” and “using foreign language for social media”. In total attitudes scores and in the sub-dimension of “emotions towards foreign language learning”, the attitudes of the participants who speak 3 languages are more positive. In the sub-dimension of “using foreign language for social media”, it was determined that the attitudes of the participants who know at least one or more languages were more positive than those who did not specify a foreign language. This result shows that knowing a foreign language encourages positive attitudes towards learning another foreign language. Based on the literature, Papuc (2016) argued that language attitudes cannot be considered separately from language differences, and that they are formed according to language diversity; he explained that the relationship between language attitudes, language diversity and linguistic discrimination is socially shaped.

The relationship between the intercultural sensitivity of Turkish origin students and their attitudes towards foreign language learning is another issue of the research. As a result of the research, it was found that as intercultural sensitivity increases, attitudes towards foreign language learning develop positively. In other words, positive attitudes can be displayed to the extent that intercultural sensitivity is internalised. Given that intercultural sensitivity is related to the development of more complex and multiple perspectives, its relationship with attitudes towards language, which is an element of culture, is an expected finding and is consistent with the results of correlational studies in the literature. Engle and Engle (2004) and Olson and Kroeger (2001) found that sensitivity follows a linear development with second language acquisition. Aydoğan and Akbarov (2014) found that attitudes towards English and the use of English have statistically significant relationships with almost all subdimensions of the intercultural sensitivity scale. Çiloğlu and Bardakçı (2019) state that the higher the language proficiency score is, the higher the intercultural sensitivity becomes. Aydoğan and Akbarov (2015) found that participants who believed they had a high level of communication abilities, skills and competencies were more sensitive and flexible.

The study also examined the effect of the sub-dimensions of the *Intercultural Sensitivity Scale* on the total scores and all the sub-dimensions of the *Attitudes Towards Foreign Language Learning Scale*. According to the results, “respect for cultural differences” was the predictor with the highest effect on the total scores of attitudes towards foreign language learning. This means that participants who accept differences with other cultures and respect their views, values, behaviours and time spent with them develop more positive attitudes towards foreign language learning. The second predictor with the highest effect is “interaction engagement”. It is understood that the participants' enjoyment of being together with different cultures, not being in a hurry to make an impression, being open-minded, giving positive reactions, not avoiding communication and showing their understanding through verbal and non-verbal messages are determinants of positive attitudes. “Interaction confidence” has the least effect on the overall scores for attitudes towards foreign language learning. Participants' reports of being confident with people from different cultures, knowing what to say, feeling comfortable and confident while speaking, and being able to interact as much as they want influenced their attitudes to be more positive. Wong (2015) states that interaction creates attitude change and increases language confidence. The fact that interaction confidence was the least effective predictor may be related to lack of experience or being influenced by prior learning/conditioning. Su (2018), who reached similar results in the opposite case of this study, examining the effect of intercultural sensitivity on attitudes towards language learning, found that attitudes towards language were a significant predictor of intercultural sensitivity in the dimensions of “interaction engagement”, “respect for cultural differences” and “interaction confidence”. On the other hand, “interaction attentiveness” was not found to be a significant predictor for total attitudes scores and almost all of the sub-dimensions. Insignificant interaction attentiveness means that during interaction with people from different cultures, the participants do not focus on the details of the messages, do not aim to obtain as much information as possible, and do not show sensitivity to unclear expressions. In addition, “interaction confidence” for the sub-dimension of “emotions towards foreign language learning”; “respect for cultural differences” for the sub-dimension of “negative thoughts towards foreign language learning”; “interaction engagement” for the sub-dimension of “using a foreign language for social media”; “interaction engagement” for the sub-dimension of “importance given to foreign language learning”; and finally “respect for cultural differences” for the sub-dimension of “thoughts towards foreign language lessons” were the predictors with the highest effect.

Like many other research, this one has some limitations. Although the countries evaluated within the scope of the research and the number of students sampled in these countries limit the generalisability of the research, it can be said that the results are statistically strong. In addition, the fact that there is no restriction as to the faculties in which the students in the study study at university can be seen as a feature that strengthens the sample from a certain point of view. The study can be repeated with a sample in which the demographic characteristics are more evenly distributed. Another limitation relates to the data collection instruments of the study. It should be taken into account that the research data are limited to the answers given to the items of two different scales, and that the *Intercultural Sensitivity Scale* is European-centred. Collecting the data with a scale that has been

prepared to be applied specifically to the Turkish world may provide more sensitivity to the results. Furthermore, the bias of reflecting what 'should be' rather than what is in quantitative self-report measures is open to criticism. More objective measures can be used in new research. Longitudinal and intervention studies can be conducted to understand the contribution of intercultural sensitivity to the individual and/or society and its reflection in communication processes. This will provide empirical evidence for the descriptive findings of this study.

The results of the study may be instructive for educational policy makers in the countries included in the study. Getting better results can be supported by the development of supporting policies. As found in this study, the resources and methods needed to develop intercultural sensitivity can be reconsidered, taking into account their impact on attitudes towards foreign language learning and thus on individual motivation. A country-specific intercultural education programme can be developed to increase the intercultural sensitivity of Turkish origin students. Intercultural sensitivity and attitudes towards foreign language learning can be investigated quantitatively or qualitatively in different geographies and samples and the predictor variables can be retested.

Declaration of Conflicting Interests

The authors declare that there is no conflict of interest.

Funding

No financial support was received from any institution in this study.

Statements of publication ethics

We hereby declare that the study has not unethical issues and that research and publication ethics have been observed carefully.

Author contribution statements

The first author designed the study and developed the theory. The second author contributed by reviewing the work and providing ideas and guidance. The first author performed all the analyses and reported the results. Both authors discussed the results and contributed to the final manuscript.

Researchers' contribution rate

The study was conducted and reported by the first author, and the second author's suggestions and guidance were utilized at every stage of the process.

Ethics Committee Approval Information

This study was discussed at the meeting numbered 08 of the Ethics Commission of Gazi University Rectorate held on 18.04.2023 and it was unanimously decided that there was no ethical objection. Research approval was notified with the letter dated 08.05.2023 and numbered 651104 (Research Code No: 2023-599).

REFERENCES

- Abaslı, K. ve Polat, Ş. (2019). Öğrencilerin kültürlerarası duyarlılık ve kültürel zekâya ilişkin görüşlerinin incelenmesi. *Anemon-Muş Alparslan Üniversitesi Sosyal Bilimler Dergisi*, 7(1), 193–202. <https://doi.org/10.18506/anemon.419526>
- Ahnagari, S. & Zamanian, J. (2014). Intercultural communicative competence in foreign language classroom. *International Journal of Academic Research in Business and Social Sciences*, 4(11), 9-16. <https://doi.org/10.6007/ijarbss/v4-i11/1265>
- Aidinlou, N.A. & Kejal, D.A. (2012). Socio-cultural factors and teaching a foreign language. *International Journal of Social Science and Education*, 2(2), 139-143.
- Alavinia, P. & Salmasi, M. (2012). On the correlation between Iranian EFL learners' shyness level and their attitudes toward language learning. *Research on Humanities and Social Sciences*, 2(9), 91-102.
- Alisaari, J. & Kilpi-Jakonen, E. (2022). Learning of intercultural competences and languages at school and their influence on global competences and immigrant-origin peers' sense of belonging. *Apples-Journal of Applied Language Studies*. 16(2), 11-40. <https://doi.org/10.47862/apples.111939>
- Altmayer, C. (2002). Kulturelle deutungsmuster in texten. Prinzipien und verfahren einer kulturwissenschaftlichen textanalyse im Fach Deutsch als Fremdsprache. *Zeitschrift für Interkulturellen Fremdsprachenunterricht*, 6(3), 1-25.
- Aydın, G. (2020). Kültür öğretiminin temel kavramları. G. Aydın (Ed.) *Yabancı/ikinci dil öğretiminde kültür ve kültürel etkileşim* içinde (pp. 1-50). Pegem Akademi.
- Aydın, G. (2021). Yabancı/ ikinci dil olarak Türkçe öğretiminde kültür ve kültürel etkileşim. D. Köksal, Ö.G. Ulum, E. Duruk (Ed.). *Türkçenin yabancı dil olarak öğretimi: Dil politikası-kültür-yöntem-teknoloji* içinde (ss. 133-156). Nobel Akademik.
- Aydoğan, H. & Akbarov, A. (2014). Attitudes and intercultural sensitivity towards English language teaching. *The International Journal of Social Sciences*, 20, 50-61.
- Aydoğan, H. & Akbarov, A. (2015). Linguo-intercultural sensitivity and its predictors in EFL in a Bosnian sample. *European Journal of Educational Research*, 4(4), 141-147. <https://doi.org/10.12973/eu-jer.4.4.141>
- Bachman, L.F. (1990). *Fundamental considerations in language testing*. Oxford University Press.

- Bayles, P. P. (2009). *Assessing the intercultural sensitivity of elementary teachers in bilingual schools in a Texas school district* (PhD Dissertation). University of Minnesota, US. (Online: https://conservancy.umn.edu/bitstream/handle/11299/49152/Bayles_umn_0130E_10245.pdf;sequence=1 / 10.03.2024).
- Bennett, J. M. (1986). A developmental approach to training for intercultural sensitivity. *International Journal of Intercultural Relations*, 10(2), 179-196. [https://doi.org/10.1016/0147-1767\(86\)90005-2](https://doi.org/10.1016/0147-1767(86)90005-2)
- Boudouaia, A., Htun, K., Al-Qadri, A., Saroh, Y. & Beddiaf, A. (2022). Intercultural sensitivity of English language teachers in Algeria. *Cogent Education* 9: 2042034, 1-13. <https://doi.org/10.1080/2331186X.2022.2042034>
- Büyüköztürk, Ş., Kılıç-Çakmak, E., Akgün, Ö., Karadeniz, Ş. & Demirel F. (2013). *Bilimsel araştırma yöntemleri* (15. Baskı). Pegem Akademi.
- Canale, M. (1983). From communicative competence to communicative pedagogy. In J.C. Richards & R.W. Schmidt (Eds.), *Language and communication* (pp. 2-27). Longman.
- Canale, M. & Swain, M. (1980). Theoretical bases of communicative approaches to secondlanguage teaching and testing. *Applied Linguistics*, 1(1), 1-47. <https://doi.org/10.1093/applin/1.1.1>
- Chen, G.-M. (1997). *A review of the concept of intercultural sensitivity*. (Online: <https://files.eric.ed.gov/fulltext/ED408634.pdf> / 10.11.2023).
- Chen, G.-M. & Starosta, W. J. (2000). The development and validation of the intercultural sensitivity scale. *Human Communication*, 3(1), 2-14.
- Chen, H. & Hu, B. (2023). On the intercultural sensitivity of university students in multicultural regions: A case study in Macao. *Frontiers in Psychology*, 14: 1090775, 1-11. <https://doi.org/10.3389/fpsyg.2023.1090775>
- Crawford-Lange, L.M. & Lange, D.L. (1987). Integrating language and culture: How to do it. *Theory Into Practice*, 26(4), 258-266. <https://doi.org/10.1080/00405848709543284>
- Çiloğlu, F. & Bardakçı, M. (2019). The Relationship between intercultural sensitivity and English language achievement. *Journal of Language and Linguistic Studies*, 15(3), 1204-1214. <https://doi.org/10.17263/jlls.631563>
- Çubukçu, F. (2013). Pre-service English teachers' intercultural sensitivity. *International Journal of Human Sciences*, 10(1), 832-843.
- Edlund A.-C., Erson E. & Milles K. (2007). *Språk och Kön*. Norstedts Akademiska Förlag.
- Ellis, R. (1994). *The study of second language acquisition*. Oxford University Press.
- Engle, L. & Engle, J. (2004). Assessing language acquisition and intercultural sensitivity development in relation to study abroad program design. *Frontiers: The Interdisciplinary Journal of Study Abroad* 10(1), 219-236. <https://doi.org/10.36366/frontiers.v10i1.142>
- Erdoğan, İ. & Okumuşlar, M. (2020). Intercultural sensitivity and ethnocentrism levels of theology students in a Turkish university sample. *Religions*, 11(5), 237. <https://doi.org/10.3390/rel11050237>
- Fakeye, D.O. (2010). Students' personal variables as correlates of academic achievement in English as a second language in Nigeria. *Journal of Social Sciences*, 22(3), 205-211. <https://doi.org/10.1080/09718923.2010.11892803>
- Fraenkel, J. R., Wallen, N. E. ve Hyun, H. H. (2012). *How to design and evaluate research in education* (8th ed.). McGraw-Hill.
- Gardner, R.C. (1985). *Social psychology and second language learning: the role of attitudes and motivation*. Edward Arnold.
- Genç, G. & Boynukara, E. (2017). Pre-Service EFL teachers' social intelligence and intercultural sensitivity. *International Journal of Language and Linguistics*, 4(4), 232-239.
- Ghamarnia, M., Soltani, A. & Rahimi, A. (2016). The interface between Iranian EFL teachers' intercultural sensitivity and their ethnic and demographic backgrounds. *Global Journal of Foreign Language Teaching*, 6(1), 36-45. <https://doi.org/10.18844/gjflt.v6i1.429>
- Hammer, M.R., Bennet, M.J. & Wiseman, R. (2003). Measuring intercultural sensitivity: The intercultural development inventory. *International Journal of Intercultural Relations*, 27(4), 421- 443. [https://doi.org/10.1016/S0147-1767\(03\)00032-4](https://doi.org/10.1016/S0147-1767(03)00032-4)
- Holm, K., Nokelainen, P. and Tirri, K. (2009). Relationship of gender and academic achievement to Finnish students' intercultural sensitivity. *High Ability Studies* 20(2), 187-200. <https://doi.org/10.1080/13598130903358543>
- Hymes, D. H. (1972). On communicative competence. In J.B. Pride & J. Holmes (Eds.), *Sociolinguistics* (pp. 269-293). Penguin Books.
- İnal, S., Evin-Gencil, İ. ve Saracaloğlu, A.S. (2005). The relation between students' attitudes toward foreign language achievement. *Dil Dergisi*, 130, 38-52. https://doi.org/10.1501/Dilder_0000000045
- İnalöz, A. B. ve Yılmaz, K. (2023). Türkçeyi yabancı dil olarak öğrenen öğrencilerin Türkçe öğrenme tutumları ile Türkçe öğrenmeye yönelik motivasyon engelleri arasındaki ilişkiler. *MANAS Sosyal Araştırmalar Dergisi*, 12(1), 62-76. <https://doi.org/10.33206/mjss.1101078>
- Jabbari, M. J. & Golkar, N. (2014). The relationships between EFL learners' language learning attitudes and language learning strategies. *International Journal of Linguistics*, 6(3), 161-167. <http://dx.doi.org/10.5296/ijl.v6i3.5837>
- Karahan, F. (2007). Language attitudes of Turkish students towards the English language and its use in Turkish context. *Çankaya University Journal of Arts and Sciences*, 1(7), 73-87.
- Karatay, H. ve Kartallıoğlu, N. (2016). Yabancı dil olarak Türkçe öğrenme tutumu ile dil becerileri edimi arasındaki ilişki. *Bolu Abant İzzet Baysal Üniversitesi Sosyal Bilimler Enstitüsü Dergisi*, 16(4), 203-214. <https://doi.org/10.11616/basbed.vi.455421>
- Karatay, H., Güngör, H. ve Çangal, Ö. (2018). Bosna Hersekli öğrencilerin Türkçe öğrenme amacı, tutumu ve öğrenme becerileri. *Kesit Akademi Dergisi*, 4 (16), 18-31.
- Kaufman, D. (2004). Constructivist issues in language learning and teaching. *Annual Review of Applied Linguistics*, 24, 303-319. <https://doi.org/10.1017/S0267190504000121>
- Koç, C. T. ve Yastıbaş, A. E. (2022). Yabancı diller yüksek okullarında eğitim gören öğrenciler için yabancı dil öğrenmeye yönelik tutum ölçeğinin geliştirilmesi. *Gazi Üniversitesi Gazi Eğitim Fakültesi Dergisi*, 42(1), 873-896. <https://doi.org/10.17152/gefad.1014605>
- Kramsch, C. (1998). *Language and culture*. Oxford University Press.
- Küllü-Sülü, A. (2014). *The role of native English speaking teachers in promoting intercultural sensitivity* (Unpublished Master's Thesis) Bilkent University, Ankara.
- Laopongharn, W. & Sercombe, P. (2009). What relevance does intercultural communication have to language education in Thailand? *ARECLS*, 6, 59- 83.
- Mellizo, J. M. (2017). Exploring intercultural sensitivity in early adolescence: A mixed methods study. *Intercultural Education*, 28(6), 571-590. <https://doi.org/10.1080/14675986.2017.1392488>
- Meydanlioğlu, A., Arkan, F. ve Gözü, S. (2015). Cultural sensitivity levels of university students receiving education in health disciplines. *Advances in Health Sciences Education*, 20(5), 1195-1204. <https://doi.org/10.1007/s10459-015-9595-z>
- Nunan, D. (1992). Sociocultural aspects of second language acquisition. *Cross Currents*, 19(1), 13-24.

- Olson, C.L. & Kroeger, K.R. (2001). Global competency and intercultural sensitivity. *Journal of Studies in International Education*, 5(2), 116–137. <https://doi.org/10.1177/102831530152003>
- Pallant, J. (2007). *SPSS survival manual: A step by step guide to data analysis using SPSS for windows* (3. Ed.). Open University Press.
- Papapavlou, A.N. (1998). Attitudes toward the Greek Cypriot dialect: sociocultural implications. *International Journal of the Sociology of Language*, 134(1), 15–28. <https://doi.org/10.1515/ijsl.1998.134.15>
- Papuc, R. D. (2016). Language variation, language attitudes and linguistic discrimination, *Essex Student Journal*, 8(1), 1–10. <https://doi.org/10.5526/esj35>
- Pavlenko, A. (2005). *Emotions and multilingualism*. Cambridge University Press.
- Penbek, Ş., Yurdakul-Şahin, D. ve Cerit, A. G. (2012). Intercultural communication competence: A study about the intercultural sensitivity of university students based on their education and international experiences. *International Journal of Logistics Systems and Management*, 11(2), 232–252. <https://doi.org/10.1504/IJLSM.2012.045425>
- Perić, B. & Radić, V. (2021). The importance of attitude in foreign language learning. *RIThink*, 10, 26–39.
- Petrovic, D. & Zlatkovic, B. (2009). Intercultural sensitivity of future primary school teachers. In Popov, at all. (Eds.) *Comparative Education, Teacher Training, Education Policy, Social Inclusion and Child Psychology*, 7, 121–128.
- Roh, S.-Z. (2014). A study on the factors affecting the intercultural sensitivity of middle and high school students in Korea. *Advanced Science and Technology Letters*, 47, 266–269. <http://dx.doi.org/10.14257/astl.2014.47.61>
- Saracaloğlu, A. S. ve Dinçer, B. (2017). The relationship between university students' attitudes towards foreign language and foreign language learning strategies. *International Journal of Language Academy*, 5(4), 440–456. <http://dx.doi.org/10.18033/ijla.3679>
- Saville-Troike, M. (2012). *Introducing second language acquisition* (2nd ed.). Cambridge University Press.
- Straffon, D. A. (2003). Assessing the intercultural sensitivity of high school students attending an international school. *International Journal of Intercultural Relations* 27(4), 487–501. [https://doi.org/10.1016/S0147-1767\(03\)00035-X](https://doi.org/10.1016/S0147-1767(03)00035-X)
- Su, Y.-C. (2018). Assessing Taiwanese college students' intercultural sensitivity, EFL interests, attitudes toward native English speakers, ethnocentrism, and their interrelation. *The Asia-Pacific Education Researcher*, 27(3), 217–226. <https://doi.org/10.1007/s40299-018-0380-7>
- Süğümlü, Ü. ve Çinpolat, E. (2023). Türkçeyi yabancı dil olarak öğrenenlerin Türkçe öğrenme kaygıları ve tutumları ile özerk öğrenme becerileri: Açımlayıcı sıralı desen çalışması. *Dokuz Eylül Üniversitesi Buca Eğitim Fakültesi Dergisi*, (58), 2859–2878. <https://doi.org/10.53444/deubefd.1331838>
- Tabachnick, B. G. & Fidell, L. S. (2013). *Using multivariate statistics* (6th ed.). Pearson Education.
- Tıkız-Ertürk, G., Demiröz, H. ve Ata, M. (2023). Intercultural competence and intercultural sensitivity levels of EFL preparatory students: Insights from a Turkish context. *Gaziantep University Journal of Social Sciences*, 22(4), 1519–1537. <https://doi.org/10.21547/jss.1310183>
- Tódor, E. M. & Dégi, Z. (2016). Language attitudes, language learning experiences and individual strategies what does school offer and what does it lack? *Acta Universitatis Sapientiae, Philologica*, 8(2), 123–137. <https://doi.org/10.1515/ausp-2016-0022>
- Turgut, M.F. ve Baykul, Y. (2021). *Eğitimde ölçme ve değerlendirme* (9. Baskı). Pegem Akademi.
- Wei, Y. (2019). Investigation into intercultural sensitivity of college non-English majors. *Frontiers in Educational Research*, 2(6), 16–22. <https://doi.org/10.25236/FER.020604>
- Wenden, A. L. (1991). *Learner strategies for learner autonomy*. Prentice Hall.
- Williams, T. R. (2005). Exploring the impact of study abroad on students' intercultural communication skills: adaptability and sensitivity. *Journal of Studies in International Education*, 9(4), 356–371. <https://doi.org/10.1177/1028315305277681>
- Wong, R. (2015). Impact of overseas immersion homestay experience on linguistic confidence and intercultural communication strategies. *Journal of Intercultural Communication*, 15(1), 1–15. <https://doi.org/10.36923/jicc.v15i1.691>
- Xin, Z. (2007). From communicative competence to communicative language teaching. *Sino-US English Teaching*, 4(9), 39–45.
- Yaprak, E. ve Uslu-Üstten, A. (2021). Yabancı dil olarak Türkçe öğrenen öğrencilerin Türkiye Türkçesine ilişkin tutumu: Gazi TÖMER örneği. *Uluslararası Türkoloji Araştırmaları ve İncelemeleri Dergisi*, 6(1), 27–36.
- Yıldırım, A. ve Şimşek, H. (2013). *Sosyal bilimlerde nitel araştırma yöntemleri*. Seçkin.
- Yılmaz, F. ve Göçen, S. (2013). Sınıf öğretmeni adaylarının kültürlerarası duyarlılık hakkındaki görüşlerinin farklı değişkenlere göre incelenmesi. *Adıyaman Üniversitesi Sosyal Bilimler Enstitüsü Dergisi*, 6(15), 374–392. <https://doi.org/10.14520/adyusbd.649>
- Yuen, C.Y.M. & Grossman, D. L. (2009). The intercultural sensitivity of student teachers in three cities. *Compare: A Journal of Comparative and International Education* 39(3), 349–65. <https://doi.org/10.1080/03057920802281571>



| Research Article / Araştırma Makalesi |

The Role of Number Sense Performance in the Relationship Between Math Anxiety and Math Achievement Among Second-Grade Primary School Students

İlkokul İkinci Sınıf Öğrencilerinde Matematik Kaygısı ile Matematik Başarısı Arasındaki İlişkide Sayı Hissi Performansının Rolü¹

Mehmet Hayri SARI², Berfin ÖZMEN³

Keywords

Primary school
Math achievement
Math anxiety
Number sense

Anahtar Kelimeler

İlkokul
Matematik başarısı
Matematik kaygısı
Sayı hissi

Received/Başvuru Tarihi

23.05.2025

Accepted / Kabul Tarihi

10.10.2025

Abstract

Purpose: This study investigates the role of number sense performance in the relationship between math anxiety and math achievement among second-grade primary school students.

Design/Methodology/Approach: The study sample consisted of 138 second-grade students from a school representing a middle socioeconomic status. Data were collected using a math anxiety scale, a math achievement test, and a number sense test. The data were analyzed through descriptive statistics, Pearson correlation, mediation, and moderation analyses.

Findings: The results revealed a strong positive correlation between number sense performance and math achievement ($r = .64$) and a moderate negative correlation between math anxiety and math achievement ($r = -.34$). Mediation analysis indicated that number sense performance partially mediated the relationship between math anxiety and math achievement. On the other hand, moderation analysis showed that number sense did not have a significant moderating effect on this relationship. Slope analysis suggested that students with high number sense performance were less affected by math anxiety.

Highlights: These findings highlight that number sense is a significant predictor of early mathematical achievement and that math anxiety indirectly influences achievement through this skill. The findings indicate that number sense-based interventions in early education may significantly enhance academic achievement by fostering both cognitive development and affective factors.

Öz

Çalışmanın amacı: Bu çalışma, ilkokul ikinci sınıf öğrencilerinde matematik kaygısı ile matematik başarısı arasındaki ilişkide sayı hissi performansının rolünü incelemektedir.

Materyal ve Yöntem: Çalışma grubunu, orta sosyoekonomik düzeydeki bir okuldan 138 ikinci sınıf öğrencisi oluşturmaktadır. Veriler, matematik kaygı ölçeği, matematik başarı testi ve sayı hissi testi aracılığıyla toplanmıştır. Veriler, betimleyici istatistik, Pearson korelasyon, aracılık ve düzenleyicilik analizleriyle değerlendirilmiştir.

Bulgular: Analizler sonucunda, sayı hissi performansı ile matematik başarısı arasında güçlü pozitif ($r = .64$), matematik kaygısı ile matematik başarısı arasında ise orta düzeyde negatif bir ilişki ($r = -.34$) bulunmuştur. Aracılık analizi, sayı hissi performansının matematik kaygısı ile matematik başarısı arasındaki ilişkide kısmi bir aracı rol oynadığını göstermiştir. Ancak düzenleyicilik analizi, sayı hissini bu ilişki üzerinde anlamlı bir düzenleyici etkisi olmadığını ortaya koymuştur. Eğitim analizi, yüksek sayı hissi performansına sahip öğrencilerin matematik kaygısından daha az etkilendiğine işaret etmektedir.

Önemli Vurgular: Elde edilen bulgular, sayı hissini, erken matematiksel başarının önemli bir yordayıcısı olduğunu ve matematik kaygısının bu beceri aracılığıyla başarıyı dolaylı olarak etkilediğini göstermektedir. Bu sonuçlar, erken yaşta yapılacak sayı hissi temelli müdahalelerin, hem bilişsel gelişimi hem de duyuşsal faktörleri destekleyerek akademik başarıyı artırmada kritik rol oynayabileceğini ortaya koymaktadır.

¹ **Corresponded Author**, Nevşehir Hacı Bektaş Veli University, Faculty of Education, Department of Basic Education, Nevşehir, TÜRKİYE; <https://orcid.org/0000-0002-7159-2635>

² Nevşehir Hacı Bektaş Veli University, Faculty of Education, Department of Basic Education, Nevşehir, TÜRKİYE; <https://orcid.org/0009-0000-5858-0675>

INTRODUCTION

Mathematical skills are increasingly acknowledged as essential for academic and professional success in today's world (Kenderov, 2022). Educational scientists, mathematics educators, neuroscientists, and psychologists have conducted extensive research over recent decades to identify the processes that influence mathematics teaching and learning. Research into the mechanisms underlying the acquisition of mathematical skills highlights a foundational construct that is believed to be innately present yet shaped and refined through experience: number sense (see Feigenson et al., 2004; Lipton & Spelke, 2003; Xue & Spelke, 2000). Before defining this core skill and examining its relationship with mathematical performance, it is equally important to highlight math anxiety (MA)—an affective factor that plays a significant role in mathematics learning and interacts with cognitive processes. This issue has received growing attention, as MA can adversely affect children's academic achievement and future employment opportunities (Ramirez et al., 2016). In light of these considerations, the present study aims to investigate the relationship between number sense performance, MA, and math achievement in young children. Understanding the interplay among these variables is crucial for developing targeted instructional interventions for students who experience difficulties with mathematical competencies and contributes to the early identification of at-risk students by highlighting key predictive factors.

Number sense performance and math achievement

In neurocognitive psychology research, number sense is defined as the ability to mentally represent and manipulate numbers and quantities (Dehaene, 2001). Within the educational paradigm (Jordan et al., 2010; National Council of Teachers of Mathematics, 1989; Reys et al., 1999), it is characterized as both the ability and disposition to develop effective strategies—applying a comprehensive understanding of numbers and operations to flexibly perform mathematical evaluations and navigate numerical situations. These mathematical skills include counting; number knowledge (comparing numerical magnitudes and distinguishing quantities); number transformations; calculations based on reference points or without them; estimations (making approximate or precise estimates by utilizing reference points); and number patterns (identifying, repeating, and extending numerical relationships) (Jordan et al., 2010; Reys et al., 1999). Although these definitions provide distinct perspectives, number sense is generally characterized by the use of diverse numerical representations; such as recognition of relative and absolute number magnitudes; selection and use of reference points; decomposition and recomposition of numbers; understanding the relational effects of operations on numbers; and the ability to perform flexible and accurate mental calculations and estimations (Reys & Yang, 1998). A consensus among neurocognitive scientists and educators highlights that number sense is strongly associated with math achievement (Caviola et al., 2020; Geary et al., 2012; Jordan et al., 2010; Mazzocco et al., 2011; Nelwan et al., 2021; Olkun, 2015; Olkun et al., 2017; Reys et al., 1999).

Previous studies have demonstrated that early number sense skills are stronger predictors of later math performance than either intelligence or working memory (e.g., Decarli et al., 2023; Friso-van den Bos et al., 2014; Mazzocco et al., 2011). For instance, a recent longitudinal study conducted by Decarli and colleagues revealed that the core number sense ability—defined as the approximate number system—measured in 60 infants at 12 months of age served as a distinct predictor of their mathematical proficiency at age four, independent of general intelligence and inhibitory control (Decarli et al., 2023). Moreover, some researchers argue that impaired number sense may underlie mathematical difficulties, suggesting that its deficiency could be a contributing factor to math-related learning challenges (e.g., Butterworth, 2010; Butterworth & Laurillard, 2010; Halberda et al., 2008; Olkun et al., 2015; Starr et al., 2013; Wilson & Dehaene, 2007). Additional evidence indicates that improvements in mathematical skills may coincide with developments in number sense, pointing to a potentially bidirectional relationship between the two constructs (e.g., Elliott et al., 2019; Mussolin et al., 2014). Conversely, other studies have questioned this association, reporting no significant correlation between number sense performance and children's mathematical achievement (e.g., Price et al., 2012; Sasanguie et al., 2012). Taken together, these findings underscore the need for further research to clarify the nature and extent of the relationship between number sense and math achievement in early childhood.

MA and math achievement

MA can be defined as a feeling of tension, worry, or even fear that interferes with performing mathematical tasks and solving mathematical problems (Ashcraft & Faust, 1994). Alternatively, it is characterized as behavioral, attitudinal, or emotional responses that negatively impact an individual's performance during the learning, application, or assessment of mathematics (McMinn & Aldridge, 2020). MA is characterized by negative emotions toward mathematics and may lead individuals to avoid math courses and math-related careers (Pizzie & Kraemer, 2017). Research has consistently shown that MA is significantly associated with math achievement (Bayırlı et al., 2021; Miller & Bichsel, 2004; Mutlu et al., 2017; Namkung et al., 2019; Sarı & Ekici, 2018; Şad et al., 2016). Importantly, MA not only undermines students' current performance but may also have long-term consequences by hindering the process of learning mathematics (Pantoja et al., 2020; Soltanlou et al., 2019; Szczygieł et al., 2024). For example, Pantoja and colleagues (2020) found that MA levels in first-grade students strongly predicted their math achievement up to third grade. Thus, MA can pose a significant obstacle to mathematical development. Regarding the directionality of the relationship between MA and math achievement, three theoretical frameworks have been proposed (Carey et al., 2016): (1) the Deficit Theory, which posits that poor performance in mathematics leads to increased anxiety; (2) the Debilitating Anxiety Model, suggesting that anxiety contributes to avoidance of learning mathematics, thereby resulting in lower achievement; and (3) the Reciprocal Theory, which views the relationship as bidirectional, where prior math achievement influences levels of MA and vice versa. Research findings indicate that both variables are likely in a vicious cycle (e.g., Foley et al., 2017; Gunderson et al., 2018;

Namkung et al., 2019; Ramirez et al., 2018; Szczygieł et al., 2024). In other words, low math achievement can lead to math anxiety, while high math anxiety can lead to low math performance.

MA typically emerges in childhood and begins to develop during the early years of primary education. As children grow older, MA levels tend to increase, generally peaking around ages 9 or 10, after which they stabilize—though symptoms may persist throughout the school years and even into adulthood (Koch, 2018). Notably, the impact of MA on children's fundamental mathematical skills appears to differ from that observed in adults and is shaped by a more intricate network of interactions (Justicia-Galiano et al., 2017). For instance, a recent meta-analysis reported a general trend of improvement in MA from primary through late secondary education, while the association between MA and math performance became more pronounced and negative during the later stages of compulsory schooling (i.e., grades 9 and 10) (Brunner et al., 2023). Similar findings have been documented in prior meta-analyses (e.g., Barroso et al., 2021; Bayırlı et al., 2021; Şad et al., 2016; Zhang et al., 2019). Although the adverse effects of MA on math performance have been observed as early as the initial school years (see Szczygieł & Pieronkiewicz, 2021), the relationship appears less consistent in children compared to adults—possibly due to the limited scope and number of previous studies focusing specifically on younger age groups.

MA and number sense performance

Although the negative effects of MA on individuals' math performance are well documented, the role of number sense performance—one of the underlying cognitive processes potentially mediating these effects—has yet to be fully clarified (Sarı & Szczygieł, 2023; Szczygieł, 2021). It is hypothesized that poor number sense skills may increase the likelihood of initial failure and negative learning experiences in mathematics education, potentially triggering MA (Lindskog et al., 2017; Maldonado Moscoso et al., 2020). Some studies examining the relationship between number sense performance and MA suggest that MA directly impairs number sense, which may contribute to lower math achievement (e.g., Ak & Ertekin, 2020; Maldonado Moscoso et al., 2020; Sarı & Szczygieł, 2023). Conversely, other studies have reported no significant correlation between number sense skills and MA (e.g., Braham & Libertus, 2018; Szczygieł, 2021). Overall, it is believed that MA may negatively influence number sense skills at early ages, thereby shaping math achievement over the long term. Moreover, the enhancement or maintenance of advanced number sense skills is considered a protective factor that may help support math performance in individuals with high levels of MA (Braham & Libertus, 2018).

In summary, when studies examining the antecedents of math achievement—particularly MA and number sense performance (Braham & Libertus, 2018; Lindskog et al., 2017; Sarı & Szczygieł, 2023; Szczygieł, 2021)—are considered collectively, it becomes evident that certain underlying mechanisms remain unclear. In particular, the relationship between MA and number sense performance, and the role this relationship plays in shaping math achievement, has yet to be fully elucidated (see Braham & Libertus, 2018; Mononen et al., 2022; Szczygieł, 2021). Consequently, there is a need for further research to better understand the mechanisms underpinning the stronger association between MA and math achievement in young children. In this context, De Smedt and colleagues (2013) argue that identifying which early cognitive and affective predictors significantly influence the development of math achievement—and understanding when and how these predictors exert their effects—can inform educational practices and curriculum design, as well as guide the development of effective early interventions.

In this context, the primary aim of the present study is to examine whether number sense performance mediates the relationship between MA and math achievement. The theoretical model tested for this purpose is presented in Figure 1. The second aim is to evaluate whether number sense performance moderates the relationship between MA and math achievement; this theoretical model is shown in Figure 2. Finally, the study seeks to determine whether number sense performance and math achievement differ between children with low and high levels of MA. Investigating differences in math performance between young children with low and high MA is important, as previous studies suggest students with higher levels of MA at later educational stages (e.g., middle school, high school, and university) may be at greater risk for poor math performance compared to those with lower levels of MA (e.g., Maloney et al., 2010; Maloney et al., 2011; Sarı & Szczygieł, 2023; Skagerlund et al., 2019).

As a result, in line with previous studies (e.g., Lindskog et al., 2017; Skagerlund et al., 2019), it is anticipated that MA will be directly associated with math achievement, and that number sense performance may partially mediate this relationship. In other words, number sense performance may help explain how MA exerts its negative effect on math achievement. Furthermore, consistent with the findings of Braham and Libertus (2018), it is expected that children with higher levels of MA and lower number sense performance will perform more poorly on math tasks than children with lower levels of MA and higher number sense performance.

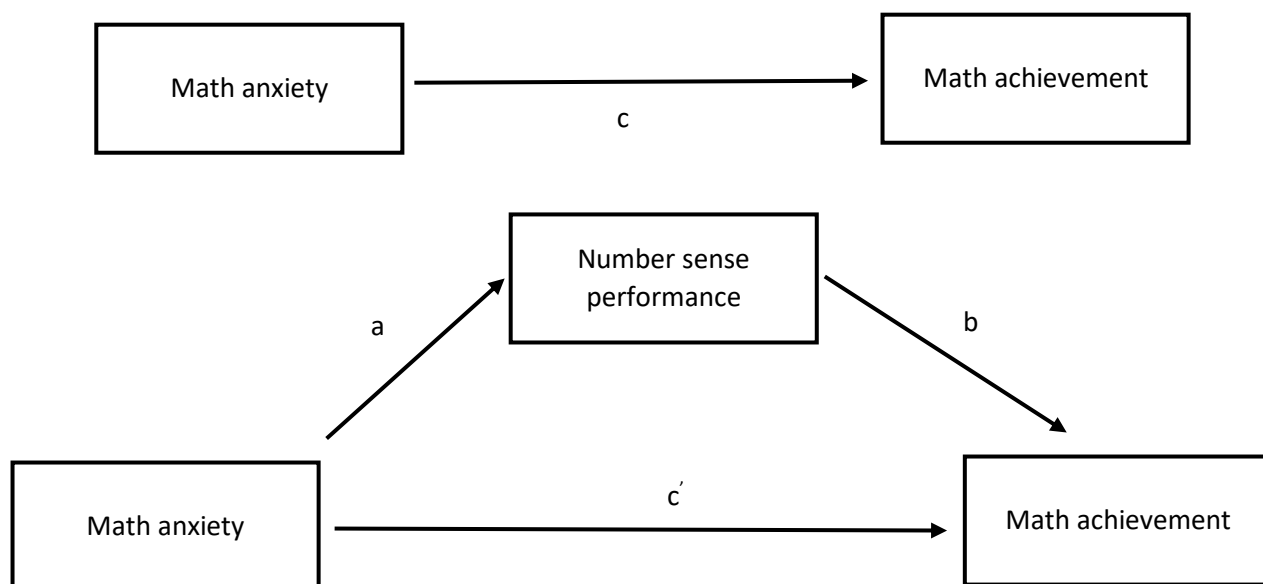


Figure 1. Mediation Model

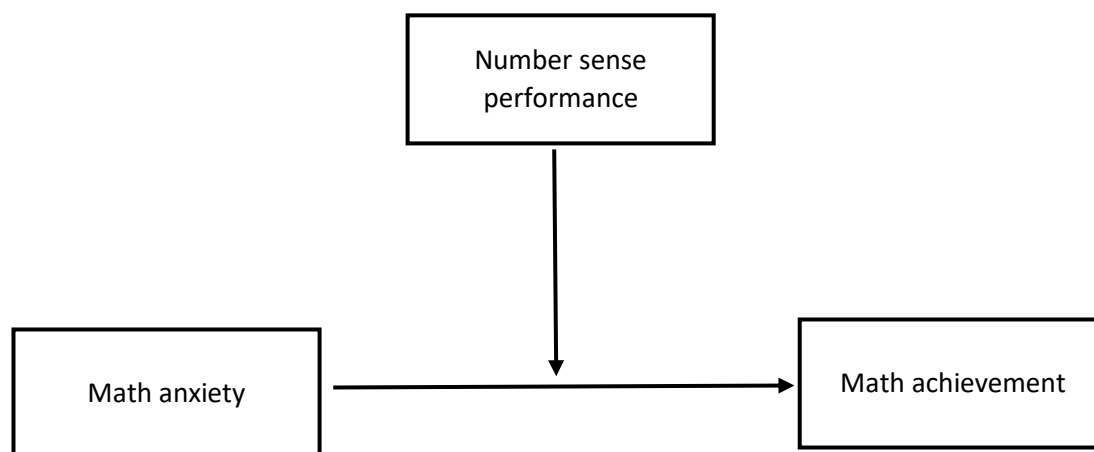


Figure 2. Moderation Model

METHOD/MATERIALS

Participants

The study sample comprised second-grade elementary school students. This age group was selected based on evidence that the early years of primary education represent a critical period for the development of fundamental mathematical skills (Wu et al., 2014) and mark the initial onset of MA (e.g., Barroso et al., 2021; Harari et al., 2013; Szczygieł & Pieronkiewicz, 2021). The minimum required sample size was calculated as 100 using the G*Power 3.1 software (Faul et al., 2009), based on the following parameters: $\alpha = 0.05$, $\beta = 0.80$, $r = -0.28$, and a two-tailed test. The effect size ($r = -0.28$) was derived from the meta-analysis conducted by Barroso et al. (2021). In total, data were collected from 141 second-grade students attending a school with a middle socioeconomic status located in the Central Anatolia Region. Following the exclusion of data from three students identified as outliers, the final sample consisted of 138 participants: 78 girls (56.5%) and 60 boys (43.5%). This sample size was deemed sufficient to ensure adequate statistical power for testing the study's hypotheses.

Data Collection Instruments

Three instruments were employed for data collection in this study: the Math Anxiety Scale, the Math Achievement Test, and the Number Sense Test.

Math Anxiety Scale: Developed by Mutlu and Söylemez (2018) to assess MA in elementary school students, this scale comprises 13 items, including 5 positively worded and 8 negatively worded items. It is designed as a three-point Likert scale, with scores

ranging from 13 (minimum) to 39 (maximum). Students were allotted one class period (40 minutes) to complete the scale. Mutlu and Söylemez (2018) reported a Cronbach's alpha (α) of 0.75 for the scale's reliability. In the present study, the Cronbach's alpha was 0.78. Confirmatory factor analysis yielded the following fit indices: $\chi^2/df = 100.739/62 = 1.63$, RMSEA = 0.068, RMR = 0.031, GFI = 0.90, CFI = 0.91, NFI = 0.81, NNFI = 0.89, and AGFI = 0.85. The χ^2/df and RMR values indicate good fit, while RMSEA, CFI, and GFI values suggest acceptable fit. Although NFI, NNFI, and AGFI values fall slightly below conventional thresholds, prior literature considers values between 0.80 and 0.89 acceptable (Cole, 1987; Doll et al., 1994; Segars & Grover, 1993). Thus, the scale's goodness-of-fit indices are deemed acceptable.

Math Achievement Test: Developed by Fidan (2013) for second-grade students, this test consists of 15 items. Students were given one class period (40 minutes) to complete the test. Correct answers were scored as 1, and incorrect answers as 0. Fidan (2013) reported a Kuder-Richardson-20 (KR-20) reliability coefficient of 0.92. In the present study, the KR-20 coefficient was 0.85.

Number Sense Test: It was developed by Palabiyik (2022) for second-graders. The test consists of 23 questions. However, since questions 20, 21, 22, and 23 contain subquestions, the subquestions in these items were evaluated separately, and the test was scored out of a total of 27 items in this study. Students were allotted one class period (40 minutes) to complete the test. Correct answers were scored as 1, and incorrect answers as 0. Palabiyik (2022) reported a KR-20 reliability coefficient of 0.86. In the present study, the KR-20 coefficient was 0.89.

Data Analysis

Prior to conducting the data analysis, the dataset was examined for outliers, as the presence of extreme values can distort the normality distribution and lead to misinterpretation of statistical results (Gürbüz, 2021). A univariate outlier analysis was performed, resulting in the identification and removal of three data points classified as outliers. Subsequently, a multivariate outlier analysis was carried out using Mahalanobis distance calculations. In accordance with the criterion proposed by Tabachnick and Fidell (2012)—where data points significant at the $p < .01$ level should be excluded—no multivariate outliers were detected in the dataset.

Following outlier removal, multicollinearity was assessed using Variance Inflation Factor (VIF) values. As recommended by O'Brien (2007), VIF values below 10 indicate no multicollinearity concerns. The computed VIF values were as follows: number sense performance (VIF = 1.72), MA (VIF = 1.14), and math achievement (VIF = 1.77). These values fall well within acceptable limits.

Means and standard deviations were calculated for all variables included in the analysis. Normality of scale and test distributions was evaluated through skewness and kurtosis indices (see Table 1). Prior to conducting the mediation and moderation analyses, Pearson product-moment correlation coefficients were computed to assess the bivariate relationships among variables (see Table 2). All descriptive and inferential analyses were performed using IBM SPSS Statistics v22. To examine whether number sense performance served as a mediating or moderating variable in the relationship between MA and math achievement, the PROCESS Macro (v4.2) was employed. Mediation was tested using Model 4 (Hayes, 2022), with 5,000 bootstrap resamples and a 95% confidence interval. The indirect effect was considered significant if the bootstrap confidence interval did not contain zero. Moderation was assessed using Model 1 (Hayes, 2022).

FINDINGS

This section presents the descriptive statistics, correlation analyses, and findings related to the mediation and moderation effects based on the data collected from second-grade students.

Descriptive Statistics

Descriptive statistics for the scores obtained from the Math Anxiety Scale, Math Achievement Test, and Number Sense Test administered to second-grade students are presented in Table 1.

Table 1. Descriptive Statistics of the Variables

	N	$\bar{X}$	SD	Skewness	Kurtosis	Min.	Max.
Number sense performance	138	9.8	6.0	0.32	-0.92	0	27
Math achievement	138	9.4	3.8	-0.61	-0.43	0	15
Math anxiety (MA)	138	18.9	4.0	0.86	0.09	13	35

Note: SD = Standard Deviation; Min. = Minimum value; Max. = Maximum value

As shown in Table 1, the mean score for number sense performance was 9.8 (SD = 6.0), with a positive skewness of 0.32 and a negative kurtosis of -0.92, indicating a left-skewed and platykurtic distribution. The mean score for math achievement was 9.4 (SD = 3.8), with a negative skewness value of -0.61, suggesting that most students achieved relatively high scores on the math

achievement test. The mean score for MA was 18.9 (SD = 4.0), with a positive skewness of 0.86. This right-skewed distribution indicates that the majority of students reported low to moderate levels of anxiety, whereas a smaller proportion of students reported high levels of anxiety. Overall, the descriptive statistics indicate that all three variables exhibit distributions that are approximately normal.

In the second phase of the study, correlational analyses were conducted to examine the relationships among the study variables. The results are presented in Table 2.

Table 2. Correlation Results Among Number Sense Performance, Math Achievement, and Math Anxiety

Variables	1	2	3
1. Number sense performance	-		
2. Math achievement	0.64**	-	
3. Math anxiety (MA)	-0.30**	-0.34**	-

Note: ** $p < .01$

The results presented in Table 2 indicate a strong positive association between number sense performance and math achievement ($r = .64, p < .01$), suggesting that students with higher fluent number sense tend to perform significantly better in mathematics. A moderate negative correlation was observed between number sense performance and MA ($r = -.30, p < .01$), implying that students with greater numerical proficiency reported lower levels of anxiety related to mathematics. Additionally, a moderate inverse relationship was found between MA and math achievement ($r = -.34, p < .01$), indicating that elevated anxiety is linked to lower performance in mathematics. Together, these findings provide preliminary support for the proposed mediation and moderation models, wherein number sense performance may play a pivotal role in explaining or altering the influence of MA on achievement.

Testing the Mediating Effect of Number Sense Performance

The results of the mediation model examining whether number sense performance (M) mediates the relationship between math anxiety (X) and math achievement (Y) are presented in Figure 3.

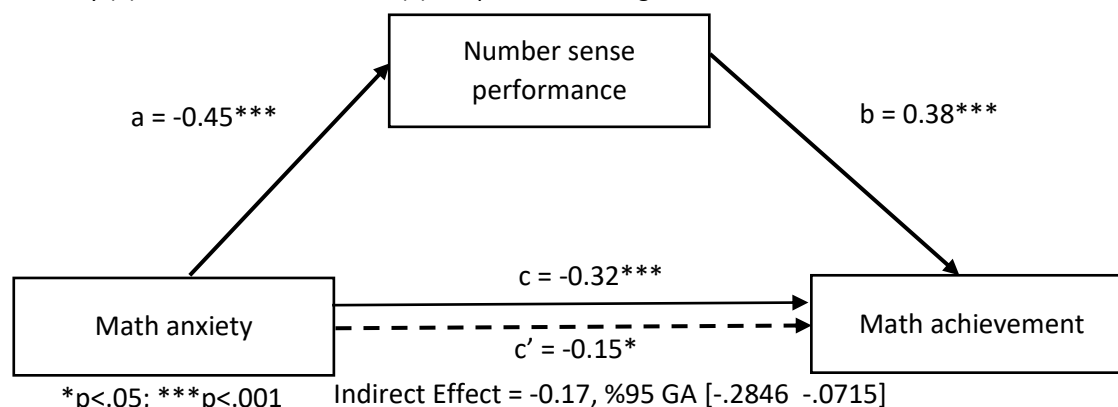


Figure 3. Mediation Model of Number Sense Performance in the Relationship Between Math Anxiety and Math Achievement

As illustrated in the mediation model presented in Figure 3, math anxiety (MA) (X) significantly predicted number sense performance (M), $a = -.45, p < .001$, and number sense performance significantly predicted math achievement (Y), $b = .38, p < .001$. The total effect of MA on math achievement was statistically significant, $c = -.32, p < .001$. These results provide evidence for a partial mediation, with both the direct effect ($c' = -.15, p < .05$) and the indirect effect ($a \times b = -.17, 95\% \text{ CI } [-.2846, -.0715]$) reaching statistical significance. The confidence interval for the indirect effect does not include zero, supporting the mediation hypothesis. In summary, higher levels of MA are associated with lower number sense performance, which in turn positively predicts math achievement. While MA has a direct negative effect on math achievement, it also exerts an indirect effect via number sense performance, thereby supporting a partial mediation model.

Testing the Moderating Effect of Number Sense Performance

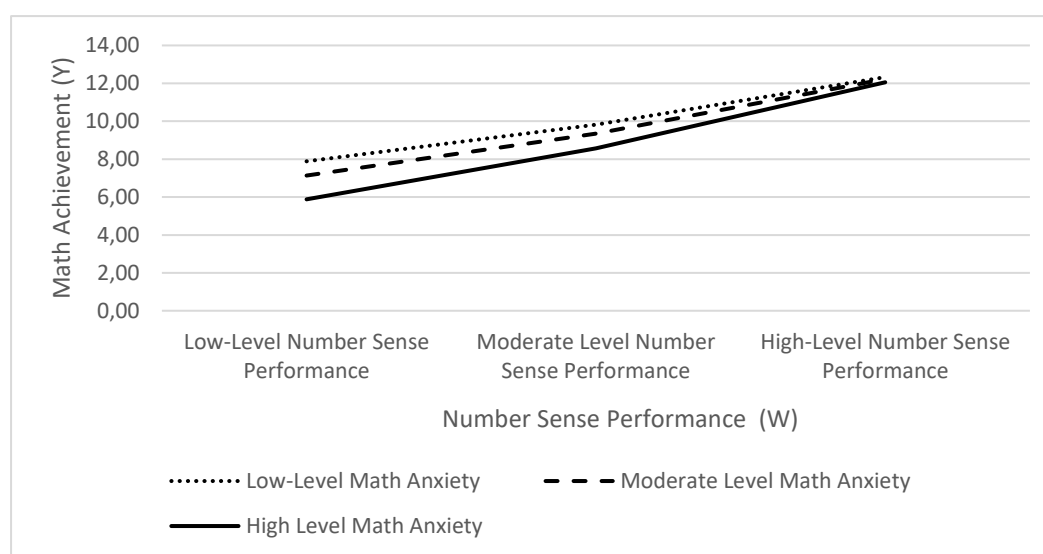
The statistical data concerning the moderation model, which examines whether number sense performance (W) moderates the relationship between math anxiety (X) and math achievement (Y), are presented in Table 3.

Table 3. Regression Analysis Results Indicating the Moderating Effect (N = 138)

Variables	β	S.E.	t	p	LLCI	ULCI
Math Anxiety (X)	-0.14	0.07	-2.20	0.00**	-0.2721	-0.0149
Number Sense Performance (W)	.038	0.04	8.82	0.03*	0.2945	0.4646
X*W (Interaction)	0.02	0.010	1.58	0.11	-0.0039	0.0353

Note: * $p < .05$; ** $p < .01$; SE: Standard Error; LLCI: Lower Limit of Confidence Interval; ULCI: Upper Limit of Confidence Interval

The results of the moderation analysis (Table 3) revealed that both math anxiety (MA) and number sense performance were significant predictors of math achievement. Specifically, MA was found to have a significant negative effect on math achievement ($\beta = -0.14$, $p = .00$), whereas number sense performance exhibited a significant positive effect ($\beta = 0.38$, $p = .03$). However, the interaction term between MA and number sense performance was not statistically significant ($\beta = 0.02$, $p = .11$), indicating that number sense performance does not moderate the relationship between MA and math achievement. In other words, the strength or direction of the relationship between MA and achievement does not significantly change at varying levels of number sense. To further investigate whether the relationship between number sense performance and math achievement varies across different levels of MA (i.e., low, moderate, and high), a simple slopes analysis was conducted. The resulting interaction patterns are illustrated in Figure 4.

**Figure 4. Simple Slope Analysis Graph of the Relationship Between Number Sense Performance and Math Achievement**

As illustrated in Figure 4, the simple slope analysis of the relationship between number sense performance and math achievement indicates that math achievement scores differ based on levels of math anxiety (MA), particularly among students with low and moderate levels of number sense performance. Specifically, in these groups, students with higher MA tend to have lower math achievement compared to their peers with moderate or low anxiety levels. In contrast, among students with high number sense performance, math achievement scores appear relatively stable and do not significantly differ across varying levels of MA. These results suggest that high number sense performance may buffer the negative effects of MA on achievement, while such a protective effect is not evident among students with lower or moderate number sense skills.

DISCUSSION, CONCLUSION AND RECOMMENDATIONS

Mathematics is a multifaceted domain encompassing a broad spectrum of cognitive subdomains. Analyzing specific mathematical competencies is crucial for advancing our understanding of affective processes—particularly MA—and its relationship to performance outcomes, as well as for exploring how task characteristics may modulate this relationship (Mielicki et al., 2022). Investigating MA during the early years of schooling not only clarifies the interplay between emotional factors and academic performance, but also informs the development of targeted interventions aimed at alleviating anxiety and enhancing achievement (Ramirez et al., 2013). In this context, the present study examines the interrelations among number sense performance, math achievement, and MA in second-grade primary school students. Specifically, it investigates whether number sense performance mediates and/or moderates the relationship between MA and math achievement, thereby contributing to a more nuanced understanding of affective–cognitive dynamics in early mathematics education.

Consistent with the proposed hypothesis, a strong and positive correlation was observed between number sense performance and math achievement ($r = .64$). This result aligns with a substantial body of research that identifies number sense as a foundational cognitive mechanism underlying mathematical competence (e.g., De Smedt et al., 2013; Geary, 2011; Jordan et al., 2010; Nelwan et al., 2021), as well as with findings from prior meta-analyses (Chen & Li, 2014; Fazio et al., 2014; Schneider et al., 2017). The strength of this association suggests that number sense plays a pivotal role in the mathematical development of second-grade students. Supporting the growth of this skill during early schooling may be crucial for fostering long-term academic success, especially given its integrative function across multiple mathematical domains.

Although number sense skills are generally thought to make a significant contribution to mathematical development, some researchers have reported evidence showing no significant correlation between these two variables (e.g., Price et al., 2012; Sasanguie et al., 2012). Inconsistencies in the literature may be related to differences in the types of tasks employed—such as symbolic versus non-symbolic comparison tasks, math achievement tests, or arithmetic performance tests (e.g., Caviola et al., 2019; Halberda et al., 2008; Nelwan et al., 2021; Olkun, 2015)—as well as to the age of the participants studied (e.g., Olkun, 2015). For instance, meta-analyses conducted by Fazio et al. (2014) and Chen and Li (2014) revealed that the correlation between number sense skills and math achievement appears strongest prior to formal schooling ($r = .40$), while this association weakens during primary education ($r = .25$) and further declines in adolescence and adulthood ($r = .22$ for ages 17 and older). These findings underscore the importance of age-related cognitive sensitivity and task design in interpreting the functional role of number sense across developmental trajectories.

Another finding from the study is that math anxiety (MA) showed moderate, negative correlations with both number sense performance ($r = -.30$) and math achievement ($r = -.34$). The observed correlations align with coefficients reported in previous cross-sectional studies (Ak & Ertekin, 2020; Sarı & Ekici, 2018; Mutlu et al., 2017) and meta-analyses (Barroso et al., 2021; Hembree, 1990; Namkung et al., 2019; Zhang et al., 2019), which have typically ranged from $r = -0.25$ to -0.40 . Similarly, meta-analyses conducted in Türkiye have reported comparable average effect sizes (e.g. Bayırlı et al., 2021; Şad et al., 2016). For instance, the meta-analysis by Bayırlı and colleagues (2021), which synthesized findings from 29 studies conducted in Türkiye, identified a significant moderate negative relationship between MA and math achievement, with an average effect size of -0.36 .

It is known that MA emerges in early childhood and tends to peak around ages 9 to 10 (Koch, 2018; Szczygieł & Pieronkiewicz, 2021). In the present study, the fact that most second-grade students reported low to moderate levels of anxiety suggests that MA may already begin to affect mathematical performance even in the early years of primary school. However, some studies have reported no significant relationship among MA, number sense, and math achievement (Braham & Libertus, 2018; Krinzinger et al., 2009; Szczygieł, 2021; Thomas & Dowker, 2000; Van den Bussche et al., 2020). These inconsistencies may be attributed to factors such as differences in the age of participants (children, adolescents, adults), the measurement tools employed (e.g., symbolic vs. non-symbolic number sense tasks; math achievement vs. arithmetic fluency tests; MA scales based on differing theoretical models), or contextual variables related to test administration (e.g., timed vs. untimed formats).

The Mediating and Moderating Role of Number Sense Performance

The most salient finding of the present study, in line with our prediction, is that number sense performance functions as a complex cognitive mechanism that not only contributes directly to math achievement but also regulates its relationship with an affective factor such as MA. In this context, the mediation analysis revealed that number sense performance plays a partial mediating role in the relationship between MA and math achievement. Specifically, while MA negatively affects number sense performance, number sense performance positively predicts math achievement. Moreover, both the direct and indirect effects were found to be statistically significant, supporting the notion that number sense partially transmits the adverse impact of anxiety on math achievement. The presence of both significant direct and indirect effects indicates that math anxiety (MA) influences achievement through two complementary pathways: (1) directly, consistent with cognitive-interference and processing-efficiency accounts that anxiety disrupts working memory and attentional control (Ashcraft & Faust, 1994; Ashcraft & Krause, 2007; Eysenck & Calvo, 1992; Eysenck et al., 2007), and (2) indirectly, aligning with evidence that MA undermines foundational number-sense processes and working-memory-dependent numerical processing (Maldonado Moscoso et al., 2020; Skagerlund et al., 2019). Furthermore, the findings are consistent with previous research suggesting that MA negatively affects core skills such as number sense (Maldonado Moscoso et al., 2020; Sarı & Szczygieł, 2023; Skagerlund et al., 2019). Although this study employed a cross-sectional design, the observed direct and indirect effects of MA on math achievement and number sense performance, as well as the direct effect of number sense on math achievement, are in line with findings from prior longitudinal studies (Cargnelutti et al., 2017; Pantoja et al., 2020; Szczygieł et al., 2024). For instance, Cargnelutti et al. (2017) demonstrated a strong negative association between MA and math achievement in second and third grades, and showed through longitudinal analyses that MA measured in second grade independently predicted math performance in third grade. Similarly, Szczygieł et al. (2024) found that math achievement at Time 1 predicted MA at Time 2, while MA at Time 1 also predicted math achievement at Time 2. Additionally, the authors noted that MA at baseline (Time 0) was predicted by general anxiety and symbolic numerical representation, an indicator of number sense skills. Collectively, these findings suggest that MA can negatively affect children's math achievement as early as second grade. Moreover, the identified mediating role of number sense underscores that anxiety can adversely impact not only math achievement but also foundational cognitive skills, highlighting the complex interplay between affective and cognitive processes in early mathematics education.

Another notable finding of the study is that the hypothesized interaction term in the moderation analysis was not statistically significant. Although number sense performance was found to partially mediate the relationship between MA and math achievement, it did not serve as a moderating factor. The lack of statistical significance in the interaction effect may be attributable to sample homogeneity or limitations inherent in the measurement instruments. However, consistent with the theoretical prediction, slope analyses revealed that students with high number sense performance appeared to be less affected by the negative impact of MA on achievement. This pattern suggests a buffering effect that goes beyond the classical moderation hypothesis. It may be posited that advanced number sense skills help students maintain their achievement levels in the presence of heightened anxiety. This interpretation is partially supported by Braham and Libertus (2018), who proposed that strong number sense may serve as a protective factor against the detrimental effects of MA. Moreover, the differentiation in math achievement scores across levels of MA (low, moderate, high) within low and moderate number sense performance groups echoes previous findings (Ashkenazi & Cohen, 2021, 2023; Maki et al., 2024; Maldonado Moscoso et al., 2020; Nunez-Pena & Suarez-Pellicioni, 2014; Passolunghi et al., 2016; Sari & Szczygiał, 2023; Skagerlund et al., 2025). These studies collectively indicate that children with high levels of MA tend to demonstrate lower performance on math tasks (Maki et al., 2024; Skagerlund et al., 2025; Sari & Szczygiał, 2023), face greater difficulty solving complex problems (Maki et al., 2024), and possess less precise representations of numerical magnitude—often making more errors in symbolic processing (Ashkenazi & Cohen, 2023; Maloney et al., 2010, 2011; Nunez-Pena & Suarez-Pellicioni, 2014; Sari & Szczygiał, 2023).

Several limitations of the present study warrant a cautious interpretation of the findings. First, the cross-sectional design restricts the ability to infer causal relationships among MA, number sense performance, and math achievement. Whether anxiety influences the development of number sense, or whether weak number sense contributes to heightened anxiety, remains unclear and can only be examined in greater detail through longitudinal or experimental research designs. Second, the exclusive use of a self-report scale to measure MA may introduce subjective bias, particularly in younger age groups where emotional awareness and self-expression skills may still be developing (see Cargnelutti et al., 2017; Primi et al., 2020). Third, the sample was drawn from a single school representing middle socioeconomic status, which may limit the generalizability of the findings to other age groups, socioeconomic backgrounds, or cultural contexts. Fourth, unmeasured variables such as teacher practices (see Hunt & Sari, 2019; Sari & Aksoy, 2016), parental attitudes toward mathematics (see Sari & Hunt, 2020; Türk & Bedir, 2021), classroom climate, and students' prior mathematical experiences (see Quintero et al., 2022) could have influenced the observed relationships. Finally, collecting data at only a single time point does not capture how these factors may evolve or interact over time. Together, these limitations suggest that the present findings should be interpreted within their specific context and highlight the need for further research to enable broader generalizations. Importantly, rather than diminishing the value of the study, these limitations offer a roadmap for future research directions aimed at deepening our understanding of the complex interplay among affective and cognitive factors in early mathematics education.

In conclusion, this study holistically examined the relationships among MA, number sense performance, and math achievement, yielding several important findings. Number sense performance emerged as a strong predictor of math achievement in second-grade primary school students. MA was found to negatively affect math achievement both directly and indirectly through number sense, which served as a partial mediating variable in this relationship. Although the moderating role of number sense was not statistically significant, slope analyses suggested that students with higher number sense performance may be less vulnerable to the adverse effects of anxiety. These results underscore the importance of designing systematic interventions aimed at enhancing number sense, particularly during the early years of primary education. Such interventions may not only support cognitive development and improve students' math achievement (e.g., Kucian et al., 2011; Sari & Olkun, 2024; Wilson et al., 2006) but also indirectly strengthen their ability to manage negative emotions related to mathematics (e.g., Ng et al., 2022; Passolunghi et al., 2020). Furthermore, comprehensive support programs that target both cognitive and affective domains could be especially beneficial for children exhibiting higher levels of MA. These programs might combine anxiety-reduction strategies with practices designed to foster number sense development. Future research should investigate the interplay between number sense and affective variables in greater depth, across different age groups and among populations at developmental risk (e.g., children with mathematical learning difficulties or those from lower socioeconomic backgrounds). Such work could provide valuable insights for early diagnosis and the prevention of mathematical difficulties. Given that the dual role of number sense—as both a mediating and protective factor in mathematical development—has been explored in only a limited number of studies (e.g., Cargnelutti et al., 2017; Pantoja et al., 2020; Szczygiał et al., 2024), there remains a need for longitudinal research to clarify these dynamics. Focusing on developmental trajectories over time may help overcome the limitations inherent in cross-sectional and correlational designs, ultimately allowing for more robust causal interpretations.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, author-ship, and/or publication of this article.

Statements of publication ethics

We hereby declare that the study has not unethical issues and that research and publication ethics have been observed carefully.

Researchers' contribution rate

Conceptualization MHS, BÖ, Data curation MHS, BÖ, Formal analysis BÖ, Investigation MHS, Methodology MHS, BÖ, Writing – original draft MHS, Writing – review & editing MHS, BÖ.

Ethics Committee Approval Information

The Research Ethics Committee of the Nevşehir HBV University (date and number = 2025/2025.05.164) authorized the recruitment, tasks, and overall procedure. The research was conducted on a voluntary basis and no reward or money was paid to the participants.

REFERENCES

- Ak, Y., & Ertekin, E. (2020). 7. ve 8. sınıf öğrencilerinin sayı hissi ile matematik kaygısı arasındaki ilişki üzerine bir çalışma. *OPUS International Journal of Society Researches*, 16(31), 4047-4076.
- Ashcraft, M. H., & Faust, M. W. (1994). Mathematics anxiety and mental arithmetic performance: An exploratory investigation. *Cognition and Emotion*, 8(2), 97-125. <https://doi.org/10.1080/02699939408408931>
- Ashcraft, M. H., & Kirk, E. P. (2001). The relationships among working memory, math anxiety, and performance. *Journal of Experimental Psychology: General*, 130(2), 224-237. <https://doi.org/10.1037/0096-3445.130.2.224>
- Ashcraft, M. H., & Krause, J. A. (2007). Working memory, math performance, and math anxiety. *Psychonomic Bulletin & Review*, 14(2), 243-248. <https://doi.org/10.3758/BF03194059>
- Ashkenazi, S., & Cohen, N. (2021). Developmental trajectories of strategy use in children with mathematical anxiety. *Acta Psychologica*, 215, 103293. <https://doi.org/10.1016/j.actpsy.2021.103293>
- Ashkenazi, S., & Cohen, N. (2023). Developmental trajectories of number line estimations in math anxiety: Evidence from bounded and unbounded number line estimation. *Applied Cognitive Psychology*, 37(6), 1316-1327. <https://doi.org/10.1002/acp.4125>
- Barroso, C., Ganley, C. M., McGraw, A. L., Geer, E. A., Hart, S. A., & Daucourt, M. C. (2021). A meta-analysis of the relation between math anxiety and math achievement. *Psychological Bulletin*, 147(2), 134. <https://doi.org/10.1037/bul0000307>
- Bayırlı, H., Geçici, M. E., & Erdem, C. (2021). Matematik kaygısı ile matematik başarıları arasındaki ilişki: Bir meta-analiz çalışması. *Pamukkale Üniversitesi Eğitim Fakültesi Dergisi*, 53, 87-109. <https://doi.org/10.9779/pauefd.783083>
- Braham, E. J., & Libertus, M. E. (2018). When approximate number acuity predicts math performance: The moderating role of math anxiety. *PLoS One*, 13(5), e0195696. <https://doi.org/10.1371/journal.pone.0195696>
- Brunner, M., Preckel, F., Götz, T., Lüdtke, O., & Keller, L. (2023). The relationship between math anxiety and math achievement: New perspectives from combining individual participant data and aggregated data in a meta-analysis. Doi: 10.31234/osf.io/6pmhr.
- Butterworth, B. (2010). Foundational numerical capacities and the origins of dyscalculia [Review]. *Trends in Cognitive Sciences*, 14(12), 534-541. <https://doi.org/10.1016/j.tics.2010.09.007>
- Butterworth, B., & Laurillard, D. (2010). Low numeracy and dyscalculia: identification and intervention. *ZDM Mathematics Education*, 42(6), 527-539. <https://doi.org/10.1007/s11858-010-0267-4>
- Carey, E., Hill, F., Devine, A., & Szűcs, D. (2016). The chicken or the egg? The direction of the relationship between mathematics anxiety and mathematics performance. *Front Psychol*, 6. <https://doi.org/10.3389/fpsyg.2015.01987>
- Cargnelutti, E., Tomasetto, C., & Passolunghi, M. C. (2017). The interplay between affective and cognitive factors in shaping early proficiency in mathematics. *Trends in Neuroscience and Education*, 8, 28-36. Doi: 10.1016/j.tine.2017.10.002
- Caviola, S., Colling, L. J., Mammarella, I. C., & Szűcs, D. (2020). Predictors of mathematics in primary school: Magnitude comparison, verbal and spatial working memory measures. *Developmental Science*, 23(6), e12957. <https://doi.org/10.1111/desc.12957>
- Caviola, S., Mammarella, I. C., & Kovas, Y. (2019). Math anxiety in children with and without mathematical difficulties: The role of gender and genetic factors. In *Mathematics anxiety: What is known and what is still to be understood*. (pp. 141-155). Routledge/Taylor & Francis Group. <https://doi.org/10.4324/9780429199981-8>
- Chen, Q., & Li, J. (2014). Association between individual differences in non-symbolic number acuity and math performance: A meta-analysis. *Acta Psychologica*, 148, 163-172. <https://doi.org/10.1016/j.actpsy.2014.01.016>
- Cole, D. A. (1987). Utility of confirmatory factor analysis in test validation research. *Journal of Consulting and Clinical Psychology*, 55(4), 584-594. <https://doi.org/10.1037/0022-006X.55.4.584>

- De Smedt, B., Noël, M.-P., Gilmore, C., & Ansari, D. (2013). How do symbolic and non-symbolic numerical magnitude processing skills relate to individual differences in children's mathematical skills? A review of evidence from brain and behavior. *Trends in Neuroscience and Education* (2), 48-55. Doi: 10.1016/j.tine.2013.06.001
- Decarli, G., Zingaro, D., Surian, L., & Piazza, M. (2023). Number sense at 12 months predicts 4-year-olds' maths skills. *Developmental Science*, 26(6), e13386. <https://doi.org/10.1111/desc.13386>
- Dehaene, S. (2001). Précis of the number sense. *Mind & Language*, 16(1), 16-36.
- Doll, W. J., Xia, W., & Torkzadeh, G. (1994). A confirmatory factor analysis of the end-user computing satisfaction instrument. *MIS Quarterly*, 18(4), 453-461. <https://doi.org/10.2307/249524>
- Elliott, L., Lisa, F., Justin, H., & Libertus, M. E. (2019). Bidirectional, Longitudinal Associations Between Math Ability and Approximate Number System Precision in Childhood. *Journal of Cognition and Development*, 20(1), 56-74. <https://doi.org/10.1080/15248372.2018.1551218>
- Eysenck, M. W., & Calvo, M. G. (1992). Anxiety and performance: The processing efficiency theory. *Cognition & Emotion*, 6(6), 409-434.
- Eysenck, M. W., Derakshan, N., Santos, R., & Calvo, M. G. (2007). Anxiety and cognitive performance: attentional control theory. *Emotion*, 7(2), 336-353.
- Faul, F., Erdfelder, E., Buchner, A., & Lang, A.-G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods*, 41(4), 1149-1160. <https://doi.org/10.3758/BRM.41.4.1149>
- Fazio, L. K., Bailey, D. H., Thompson, C. A., & Siegler, R. S. (2014). Relations of different types of numerical magnitude representations to each other and to mathematics achievement. *Journal of Experimental Child Psychology*, 123, 53-72.
- Feigenson, L., Dehaene, S., & Spelke, E. (2004). Core systems of number. *Trends in Cognitive Sciences*, 8(7), 307-314. <https://doi.org/10.1016/j.tics.2004.05.002>
- Fidan, E. (2013). *İlkokul öğrencileri için matematik dersi sayılar öğrenme alanında başarı testi geliştirilmesi* [Yüksek lisans tezi, Ankara Üniversitesi]. YÖK Ulusal Tez Merkezi.
- Foley, A. E., Herts, J. B., Borgonovi, F., Guerriero, S., Levine, S. C., & Beilock, S. L. (2017). The Math Anxiety-Performance Link: A Global Phenomenon. *Current Directions in Psychological Science*, 26(1), 52-58. <https://doi.org/10.1177/0963721416672463>
- Friso-van den Bos, I., Kroesbergen, E. H., & Van Luit, J. E. (2014). Number sense in kindergarten children: Factor structure and working memory predictors. *Learning and Individual Differences*, 33, 23-29.
- Geary, D. C. (2011). Cognitive predictors of achievement growth in mathematics: a 5-year longitudinal study. *Dev. Psychol.*, 47. <https://doi.org/10.1037/a0025510>
- Geary, D. C., Hoard, M. K., & Bailey, D. H. (2012). Fact retrieval deficits in low achieving children and children with mathematical learning disability. *Journal of Learning Disabilities*, 45(4), 291-307. <https://doi.org/10.1177/0022219410392046>
- Gunderson, E. A., Park, D., Maloney, E. A., Beilock, S. L., & Levine, S. C. (2018). Reciprocal relations among motivational frameworks, math anxiety, and math achievement in early elementary school. *Journal of Cognition and Development*, 19(1), 21-46. <https://doi.org/10.1080/15248372.2017.1421538>
- Gürbüz, S. (2021). *Sosyal bilimlerde aracı, düzenleyici ve durumsal etki analizleri* (2. basım). Seçkin Yayıncılık.
- Halberda, J., Mazocco, M. M., & Feigenson, L. (2008). Individual differences in non-verbal number acuity correlate with maths achievement. *Nature*, 455(7213), 665-668. <https://doi.org/10.1038/nature07246>
- Harari, R. R., Vukovic, R. K., & Bailey, S. P. (2013). Mathematics anxiety in young children: An exploratory study. *The Journal of Experimental Education*, 81(4), 538-555. <https://doi.org/10.1080/00220973.2012.727888>
- Hayes, A. F. (2022). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach* (3. rd ed.). The Guilford Press.
- Hembree, R. (1990). The nature, effects, and relief of mathematics anxiety. *Journal for Research in Mathematics Education*, 21(1), 33-46. <https://doi.org/10.2307/749455>
- Hunt, T. E., & Sari, M. H. (2019). An English version of the Mathematics Teaching Anxiety Scale. *International Journal of Assessment Tools in Education*, 6(3), 436-443. <https://doi.org/10.21449/ijate.615640>
- Jordan, N. C., Glutting, J., & Ramineni, C. (2010). The importance of number sense to mathematics achievement in first and third grades. *Learning and Individual Differences*, 20(2), 82-88. <https://doi.org/10.1016/j.lindif.2009.07.004>
- Justicia-Galiano, M. J., Martín-Puga, M. E., Linares, R., & Pelegrina, S. (2017). Math anxiety and math performance in children: The mediating roles of working memory and math self-concept. *British Journal of Educational Psychology*, 87(4), 573-589. <https://doi.org/10.1111/bjep.12165>
- Kenderov, P. S. (2022). Mathematics competitions: an integral part of the educational process. *ZDM Mathematics Education* 54, 983-996. Doi: 10.1007/s11858-022-01348-4

- Koch, I. (2018). Maths anxiety: Students, pre-and in-service teachers. *AMSI Choose Maths Research*, 4, 1-32. https://amsi.org.au/wp-content/uploads/2019/01/researchreport4-maths_anxiety_students_and_teachers.pdf
- Krinzinger, H., Kaufmann, L., & Willmes, K. (2009). Math anxiety and math ability in early primary school years. *Journal Psychoeduc Assess*, 27(3), 206-225. <https://doi.org/10.1177/0734282908330583>
- Kucian, K., Grond, U., Rotzer, S., Henzi, B., Schonmann, C., Plangger, F., Galli, M., Martin, E., & von Aster, M. (2011). Mental number line training in children with developmental dyscalculia [Research Support, Non-U.S. Gov't]. *Neuroimage*, 57(3), 782-795. <https://doi.org/10.1016/j.neuroimage.2011.01.070>
- Lindskog, M., Winman, A., & Poom, L. (2017). Individual differences in nonverbal number skills predict math anxiety. *Cognition*, 159, 156-162. <https://doi.org/10.1016/j.cognition.2016.11.014>
- Lipton, J., & Spelke, E. (2003). Origins of number sense: Large-number discrimination in human infants. *Psychological Science*, 14(5), 396-401.
- Maki, K. E., Zaslofsky, A. F., Coddling, R., & Woods, B. (2024). Math anxiety in elementary students: Examining the role of timing and task complexity. *Journal of School Psychology*, 106, 101316. <https://doi.org/10.1016/j.jsp.2024.101316>
- Maldonado Moscoso, P. A., Anobile, G., Primi, C., & Arrighi, R. (2020). Math anxiety mediates the link between number sense and math achievements in high math anxiety young adults. *Front Psychol*, 11. <https://doi.org/10.3389/fpsyg.2020.01095>
- Maloney, E. A., Ansari, D., & Fugelsang, J. A. (2011). Rapid communication: the effect of mathematics anxiety on the processing of numerical magnitude. *Q J Exper Psychol*, 64(1), 10-16. <https://doi.org/10.1080/17470218.2010.533278>
- Maloney, E. A., & Beilock, S. L. (2012). Math anxiety: Who has it, why it develops, and how to guard against it. *Trends in Cognitive Sciences*, 16(8), 404-406.
- Maloney, E. A., Risko, E. F., Ansari, D., & Fugelsang, J. (2010). Mathematics anxiety affects counting but not subitizing during visual enumeration. *Cognition*, 114(2), 293-297. <https://doi.org/10.1016/j.cognition.2009.09.013>
- Mazzocco, M. M., Feigenson, L., & Halberda, J. (2011). Impaired acuity of the approximate number system underlies mathematical learning disability (dyscalculia). *Child Development*, 82(4), 1224-1237. <https://doi.org/10.1111/j.1467-8624.2011.01608.x>
- McMinn, M., & Aldridge, J. (2020). Learning environment and anxiety for learning and teaching mathematics among preservice teachers. *Learning Environments Research*, 23(3), 331-345. <https://doi.org/10.1007/s10984-019-09304-y>
- Mielicki, M. K., Wilkey, E. D., Scheibe, D. A., Fitzsimmons, C. J., Sidney, P. G., Bellon, E., Ribner, A. D., Soltanlou, M., Starling-Alves, I., Coolen, I., Ansari, D., & Thompson, C. A. (2023). Task features change the relation between math anxiety and number line estimation performance with rational numbers: Two large-scale online studies. *Journal of Experimental Psychology: General*, 152(7), 2094-2117. <https://doi.org/10.1037/xge0001382>
- Miller, H., & Bichsel, J. (2004). Anxiety, working memory, gender, and math performance. *Personality and Individual Differences*, 37(3), 591-606. <https://doi.org/10.1016/j.paid.2003.09.029>
- Mononen, R., Niemivirta, M., Korhonen, J., Lindskog, M., & Tapola, A. (2022). Developmental relations between mathematics anxiety, symbolic numerical magnitude processing and arithmetic skills from first to second grade. *Cognition and Emotion*, 36(3), 452-472. <https://doi.org/10.1080/02699931.2021.2015296>
- Mussolin, C., Nys, J., Content, A., & Leybaert, J. (2014). Symbolic number abilities predict later approximate number system acuity in preschool children. *PLoS One*, 9(3), e91839. <https://doi.org/10.1371/journal.pone.0091839>
- Mutlu, Y., & Söylemez, İ. (2018). İlkokul 3. ve 4. sınıf çocukları için Matematik Kaygı Ölçeği; güvenirlik ve geçerlik çalışması. *EKEV Akademi Dergisi*, (73), 429-440.
- Mutlu, Y., Söylemez, İ., & Yasul, A. F. (2017). İlkokul öğrencilerinin matematik kaygısı ile matematik başarıları arasındaki ilişkinin incelenmesi. *Journal of Human Sciences*, 14(4), 4425-4434. doi:10.14687/jhs.v14i4.5019
- National Council of Teachers of Mathematics (1989). *Curriculum and evaluation standards for school mathematics*. National Council of Teachers of Mathematics Reston, Va.
- Namkung, J. M., Peng, P., & Lin, X. (2019). The relation between mathematics anxiety and mathematics performance among school-aged students: a meta-analysis. *Review of Educational Research*, 89(3), 459-496. Doi: 10.3102/0034654319843494
- Nelwan, M., Friso-van den Bos, I., Vissers, C., & Kroesbergen, E. (2021). The relation between working memory, number sense, and mathematics throughout primary education in children with and without mathematical difficulties. *Child Neuropsychology*, 28(2), 143-170. Doi: 10.1080/09297049.2021.1959905
- Ng, C.-T., Chen, Y.-H., Wu, C.-J., & Chang, T.-T. (2022). Evaluation of math anxiety and its remediation through a digital training program in mathematics for first and second graders. *Brain and Behavior*, 12, e2557. <https://doi.org/10.1002/brb3.2557>
- Nunez-Pena, M. I., & Suarez-Pellicioni, M. (2014). Less precise representation of numerical magnitude in high math-anxious individuals: an ERP study of the size and distance effects. *Biol Psychol*, 103, 176-183. <https://doi.org/10.1016/j.biopsycho.2014.09.004>
- O'Brien, R. M. (2007). A caution regarding rules of thumb for variance inflation factors. *Quality & Quantity*, 41(5), 673-690. <https://doi.org/10.1007/s11135-006-9018-6>

- Olkun, S. (2015). *6-11 Yaş Türk çocukları örnekleminde diskalkuliye yatkınlığı ayırt etmede kullanılacak bir ölçme aracı geliştirme çalışması*. 111k545 nolu TÜBİTAK projesi. Ankara, Türkiye.
- Olkun, S., Altun, A., Göçer Şahin, S., & Akkurt Denizli, Z. (2015). Deficits in basic number competencies may cause low numeracy in primary school children. *Education and Science*, 40(177), 141–159. <https://doi.org/10.15390/eb.2015.3287>
- Olkun, S., Mutlu, Y., & Sarı, M. H. (2017, 24-26 June). *The relationships between number sense and mathematics achievement*. International Conference on Education and New Developments, Lisbon, Portugal.
- Palabıyık, E. (2022). *Okul öncesi ve ilkökul öğrencilerinin sayı hissi becerilerinin incelenmesi* [Doktora tezi, Gazi Üniversitesi]. YÖK Ulusal Tez Merkezi.
- Pantoja, N., Schaeffer, M. W., Rozek, C. S., Beilock, S. L., & Levine, S. C. (2020). Children's math anxiety predicts their math achievement over and above a key foundational math skill. *Journal of Cognition and Development*, 21(5), 709-728. <https://doi.org/10.1080/15248372.2020.1832098>
- Passolunghi, M. C., Caviola, S., De Agostini, R., Perin, C., & Mammarella, I. C. (2016). Mathematics anxiety, working memory, and mathematics performance in secondary-school children [Original Research]. *Frontiers in Psychology*, 7(42), 1-8. <https://doi.org/10.3389/fpsyg.2016.00042>
- Passolunghi, M. C., De Vita, C., & Pellizzoni, S. (2020). Math anxiety and math achievement: The effects of emotional and math strategy training. *Developmental Science*, 23(6), e12964. <https://doi.org/10.1111/desc.12964>
- Pizzie, R. G., & Kraemer, D. J. M. (2017). Avoiding math on a rapid timescale: Emotional responsivity and anxious attention in math anxiety. *Brain and Cognition*, 118, 100-107. <https://doi.org/10.1016/j.bandc.2017.08.004>
- Price, G. R., Palmer, D., Battista, C., & Ansari, D. (2012). Nonsymbolic numerical magnitude comparison: Reliability and validity of different task variants and outcome measures, and their relationship to arithmetic achievement in adults. *Acta Psychologica* 140, 50-57. <https://doi.org/10.1016/j.actpsy.2012.02.008>
- Primi, C., Donati, M. A., Izzo, V. A., Guardabassi, V., O'Connor, P. A., Tomasetto, C., & Morsanyi, K. (2020). The Early Elementary School Abbreviated Math Anxiety Scale (the EES-AMAS): A New Adapted Version of the AMAS to Measure Math Anxiety in Young Children [Original Research]. *Frontiers in Psychology*, 11. <https://doi.org/10.3389/fpsyg.2020.01014>
- Quintero, M., Hasty, L., Li, T., Song, S., & Wang, Z. (2022). A multidimensional examination of math anxiety and engagement on math achievement. *British Journal of Educational Psychology*, 92(3), 955-973. <https://doi.org/10.1111/bjep.12482>
- Ramirez, G., Chang, H., Maloney, E. A., Levine, S. C., & Beilock, S. L. (2016). On the relationship between math anxiety and math achievement in early elementary school: The role of problem solving strategies. *Journal of Experimental Child Psychology*, 141, 83-100. <https://doi.org/10.1016/j.jecp.2015.07.014>
- Ramirez, G., Gunderson, E. A., Levine, S. C., & Beilock, S. L. (2013). Math anxiety, working memory, and math achievement in early elementary school. *Journal of Cognition and Development*, 14(2), 187-202. Doi: 10.1080/15248372.2012.664593
- Ramirez, G., Shaw, S. T., & Maloney, E. A. (2018). Math anxiety: Past research, promising interventions, and a new interpretation framework. *Educational Psychologist*, 53(3), 145-164.
- Reys, R., Reys, B., Emanuelsson, G., Johansson, B., McIntosh, A., & Yang, D. C. (1999). Assessing Number Sense of Students in Australia, Sweden, Taiwan, and the United States. *School Science and Mathematics*, 99(2), 61-70. <https://doi.org/10.1111/j.1949-8594.1999.tb17449.x>
- Reys, R. E., & Yang, D.-C. (1998). Relationship between computational performance and number sense among sixth- and eighth-grade students in Taiwan. *Journal for Research in Mathematics Education JRME*, 29(2), 225-237. <https://doi.org/10.5951/jresmetheduc.29.2.0225>
- Sarı, M. H., & Aksoy, N. C. (2016). Sınıf öğretmenlerinin matematik öğretimi kaygısı ile öğretme stilleri tercihleri arasındaki ilişki. *Turkish Studies*, 11(3), 1953-1968.
- Sarı, M. H., & Ekici, G. (2018). İlkokul 4. sınıf öğrencilerinin matematik başarıları ile aritmetik performanslarını etkileyen duyuşsal değişkenlerin belirlenmesi. *OPUS International Journal of Society Researches*, 8(15), 1562-1594. <https://doi.org/10.26466/opus.451025>
- Sarı, M. H., & Hunt, T. E. (2020). Parent-child mathematics affect as predictors of children's mathematics achievement. *International Online Journal of Primary Education*, 9(1), 85-96.
- Sarı, M. H., & Olkun, S. (2024). Improving low mathematics achievers' number sense via number line training with board games. *Journal of Education and Future*, 26, 41-56. <https://doi.org/10.30786/jef.1371037>
- Sarı, M. H., & Szczygiel, M. (2023). The role of math anxiety in the relationship between approximate number system and math performance in young children. *Psychology in the Schools*, 60(4), 912-930. Doi: 10.1002/pits.22794
- Sasanguie, D., De Smedt, B., Defever, E., & Reynvoet, B. (2012). Association between basic numerical abilities and mathematics achievement. *British Journal of Developmental Psychology*, 30(2), 344-357. Doi: 10.1111/j.2044-835X.2011.02048.x
- Segars, A. H., & Grover, V. (1993). Re-examining perceived ease of use and usefulness: A confirmatory factor analysis. *MIS Quarterly*, 17(4), 517–525. Doi: 10.2307/249590

- Schneider, M., Beeres, K., Coban, L., Merz, S., Schmidt, S. S., Stricker, J., & Smedt, B. (2017). Associations of non-symbolic and symbolic numerical magnitude processing with mathematical competence: a meta-analysis. *Developmental Science*, 20(3), e12372. <https://doi.org/10.1111/desc.12372>
- Skagerlund, K., Östergren, R., Västfjäll, D., & Träff, U. (2019). How does mathematics anxiety impair mathematical abilities? Investigating the link between math anxiety, working memory, and number processing. *PLoS One*, 14(1), e0211283. <https://doi.org/10.1371/journal.pone.0211283>
- Skagerlund, K., Skagenholt, M., & Träff, U. (2025). Mathematics anxiety and number processing: The link between executive functions, cardinality, and ordinality. *Quarterly Journal of Experimental Psychology*, 78(1), 101-115. <https://doi.org/10.1177/17470218241234041>
- Soltanlou, M., Artemenko, C., Dresler, T., Fallgatter, A. J., Ehlis, A.-C., & Nuerk, H.-C. (2019). Math anxiety in combination with low visuospatial memory impairs math learning in children [Original Research]. *Frontiers in Psychology*, 10(89). <https://doi.org/10.3389/fpsyg.2019.00089>
- Starr, A., Libertus, M. E., & Brannon, E. M. (2013). Number sense in infancy predicts mathematical abilities in childhood. *Proceedings of the National Academy of Sciences*, 110(45), 18116-18120.
- Szczygieł, M. (2021). The relationship between math anxiety and math achievement in young children is mediated through working memory, not by number sense, and it is not direct. *Contemporary Educational Psychology*, 65, 101949. <https://doi.org/10.1016/j.cedpsych.2021.101949>
- Szczygieł, M., & Pieronkiewicz, B. (2021). Exploring the nature of math anxiety in young children: Intensity, prevalence, reasons. *Mathematical Thinking and Learning*, 24(3), 248-266. <https://doi.org/10.1080/10986065.2021.1882363>
- Szczygieł, M., Szűcs, D., & Toffalini, E. (2024). Math anxiety and math achievement in primary school children: Longitudinal relationship and predictors. *Learning and Instruction*, 92, 101906. <https://doi.org/10.1016/j.learninstruc.2024.101906>
- Şad, S. N., Kış, A., Demir, M., & Özer, N. (2016). Matematik başarısı ile matematik kaygısı arasındaki ilişki üzerine bir meta-analiz çalışması. *Pegem Eğitim ve Öğretim Dergisi*, 6(3), 371-392. <http://dx.doi.org/10.14527/pegegog.2016.019>
- Tabachnick, B. G., & Fidell, L. S. (2013). *Using multivariate statistics* (6th ed.). Pearson.
- Thomas, G., & Dowker, A. (2000). *Mathematics anxiety and related factors in young children*. British Psychological Society Developmental Section Conference,
- Türk, C., & Bedir, G. (2021). İlkokul 4. sınıf öğrencilerinin matematik kaygısı ile ders başarısı arasındaki ilişki. *Journal of Educational Reflections*, 5(2), 73-88.
- Van den Bussche, E., Vanmeert, K., Aben, B., & Sasanguie, D. (2020). Too anxious to control: the relation between math anxiety and inhibitory control processes. *Scientific Reports*, 10(1), 19922. <https://doi.org/10.1038/s41598-020-76920-7>
- Wilson, A. J., & Dehaene, S. (2007). Number sense and developmental dyscalculia. In D. Coch, G. Dawson, & K. Fischer (Eds.), *human behavior, learning, and the developing brain: atypical development*. Guilford Press.
- Wilson, A. J., Revkin, S. K., Cohen, D., Cohen, L., & Dehaene, S. (2006). An open trial assessment of "The Number Race", an adaptive computer game for remediation of dyscalculia. *Behavioral and Brain Functions*, 2(1), 20. <https://doi.org/10.1186/1744-9081-2-20>
- Wu, S. S., Willcutt, E. G., Escovar, E., & Menon, V. (2014). Mathematics achievement and anxiety and their relation to internalizing and externalizing behaviors. *Journal of Learning Disabilities*, 47(6), 503-514. <https://doi.org/10.1177/0022219412473154>
- Xue, F., & Spelke, E. (2000). Large number discrimination in 6-month-old infants. *Cognition* 74, B1-B11.
- Zhang, J., Zhao, N., & Kong, Q. P. (2019). The relationship between math anxiety and math performance: A meta-analytic investigation [Systematic Review]. *Frontiers in Psychology*, 10. <https://doi.org/10.3389/fpsyg.2019.01613>



| Research Article / Araştırma Makalesi |

A Psychometric Analysis of The Mathematics Learning Disability Screening Scale Developed Through DSM-5

DSM-5 Aracılığıyla Geliştirilen Matematik Öğrenme Güçlüğü Tarama Ölçeğinin Psikometrik Analizi

Tunahan FİLİZ¹, Ahmet ÇELİK²

Keywords

1. Mathematical learning disability
2. Scale development
3. Validity
4. Reliability

Anahtar Kelimeler

1. Matematik öğrenme güçlüğü
2. Ölçek Geliştirme
3. Güvenirlik
4. Geçerlik

Received/Başvuru Tarihi
20.03.2025

Accepted / Kabul Tarihi
10.10.2025

Abstract

Purpose: This study, it was aimed to develop a valid and reliable measurement tool by evaluating the psychometric properties of the Mathematics Learning Disabilities Screening Scale (MLDSS) for elementary and secondary school students. In this context, a measurement tool that is culturally and linguistically appropriate for the Turkish context and overcomes the limitations of translated scales was developed.

Design/Methodology/Approach: The study was conducted as a scale development study within a survey research design. The scale's items were developed based on DSM-5 criteria and an extensive literature review. The sample consisted of 644 students, identified by their teachers, from 120 schools across Türkiye's seven geographical regions. The psychometric properties of the scale were evaluated using Exploratory Factor Analysis (EFA) and Confirmatory Factor Analysis (CFA), along with internal consistency (Cronbach's Alpha) and test-retest reliability analyses.

Findings: The EFA results revealed a three-factor structure (Number Sense, Calculation, and Mathematical Reasoning) that explained 68.5% of the total variance. The CFA confirmed this structure, with goodness-of-fit indices indicating an excellent model fit (e.g., CFI = .95, IFI = .95, RMSEA = .069, $\chi^2/df = 2.41$). The scale demonstrated high internal consistency (Cronbach's Alpha = .93) and strong test-retest reliability ($r = .90$).

Highlights: The mathematics learning disability screening scale is a valid and reliable instrument developed for elementary and secondary school students. Its three-factor structure is consistent with modern theories of mathematics learning disabilities and its strong psychometric properties make it a valuable tool for educators and experts in early identification and intervention planning.

Öz

Çalışmanın amacı: Bu çalışmada, ilkökul ve ortaokul öğrencileri için Matematik Öğrenme Güçlüğü Tarama Ölçeği'nin (MLDSS) psikometrik özellikleri değerlendirilerek geçerli ve güvenilir bir ölçeğin geliştirilmesi amaçlanmıştır. Bu kapsamda, Türkiye bağlamına kültürel ve dilsel olarak uygun, çeviri ölçeklerin sınırlılıklarını gideren bir ölçeğin geliştirilmiştir.

Materyal ve Yöntem: Araştırma, tarama modeline uygun bir ölçek geliştirme çalışması olarak yürütülmüştür. Ölçek maddeleri DSM-5 kriterleri ve kapsamlı literatür taramasına dayalı olarak geliştirilmiştir. Örnekleme, Türkiye'nin yedi coğrafi bölgesindeki 120 okuldan, öğretmenleri tarafından belirlenen 644 öğrenciden oluşmaktadır. Ölçeğin psikometrik özellikleri, Açıklayıcı Faktör Analizi (AFA), Doğrulayıcı Faktör Analizi (DFA), iç tutarlılık (Cronbach Alfa) ve test-tekrar test güvenirlik analizleri kullanılarak değerlendirilmiştir.

Bulgular: AFA sonuçları, toplam varyansın %68.5'ini açıklayan üç faktörlü bir yapı (Sayı Hissi, Hesaplama ve Matematiksel Akıl Yürütme) ortaya koymuştur. DFA, bu yapıyı doğrulamış ve uyum indeksleri mükemmel bir model uyumuna işaret etmiştir (örneğin, CFI = .95, IFI = .95, RMSEA = .069, $\chi^2/df = 2.41$). Ölçek, yüksek düzeyde iç tutarlılık (Cronbach Alfa = .93) ve güçlü test-tekrar test güvenirliği ($r = .90$) göstermiştir.

Önemli Vurgular: Matematik öğrenme güçlüğü tarama ölçeği, ilkökul ve ortaokul öğrencileri için geliştirilmiş geçerli ve güvenilir bir araçtır. Üç faktörlü yapısı, modern matematik öğrenme güçlüğü teorileriyle uyumlu olup güçlü psikometrik özellikleri sayesinde eğitimciler ve uzmanlar için erken tanılama ve müdahale planlamasında değerli bir araç olma potansiyeli taşımaktadır.

¹ Corresponded Author, Bayburt University, Faculty of Education, Department of Educational Sciences, Bayburt, Türkiye; <https://orcid.org/0000-0002-3149-8783>

² Giresun Provincial Directorate of National Education, Giresun, Türkiye; <https://orcid.org/0000-0001-5145-7041>

INTRODUCTION

A mathematical learning disability (MLD) is characterized by persistent and significant difficulties in comprehending mathematical concepts, applying operations, and solving problems. These challenges are not related to an individual's overall cognitive ability but are typically caused by disruptions in specific cognitive functions (Asfuroğlu & Fidan, 2016; Geary, 2004). Consequently, students with MLD may struggle with understanding mathematical ideas, performing calculations accurately and efficiently, and applying mathematical knowledge to real-world situations (Akin & Sezer, 2010; Deruaz et al., 2020). These persistent difficulties can undermine a student's self-confidence, heighten anxiety, and negatively impact their motivation, making early and accurate identification crucial for providing effective educational interventions (Mutlu, 2019; 2024).

This study was designed to remedy the notable lack of culturally appropriate assessment tools for Mathematical Learning Disability (MLD) in Türkiye. Although many international scales exist, a heavy reliance on direct translations of these foreign instruments has led to critical concerns about their cross-cultural validity. It is questionable whether such tools can accurately measure the same constructs in Turkish students. In response to this specific gap, the present research focused on developing a new Turkish MLD rating scale. The primary objective was to create and psychometrically validate an assessment tool, tailored to the local context, that can effectively and reliably identify primary and secondary school students who may have MLD.

Theoretical Framework

Mathematical Learning Disability

MLD is a specific learning disorder defined as a structural deficit in "number sense"—an innate neurological system that supports an individual's basic skills in perceiving, comparing, and using numerical quantities (Butterworth, 2010; Von Aster & Shalev, 2007). This core deficit leads to a heterogeneous profile, manifesting differently both cognitively and academically among individuals. MLD can cause various problems, extending beyond understanding abstract mathematical concepts to include challenges in remembering and applying procedures (Cirino et al., 2015; Geary, 2013; Olkun, 2015; Filiz & Güneş, 2023). This diversity highlights the critical importance of early diagnosis. Interventions are particularly effective at a young age when the brain is highly malleable, as they can more effectively shape the learning process and increase the likelihood of significant improvements in students' mathematical skills (Mutlu et al., 2021; Uçar & Duman, 2024; Tosto et al., 2015). Despite these challenges, students with MLD can learn when supported by appropriate instructional interventions, for which an accurate diagnosis is a prerequisite. The development of a reliable diagnostic tool, as undertaken in this study, is therefore essential for facilitating these effective interventions.

The varied nature of Mathematical Learning Disability (MLD) is reflected in its diverse definitions within the literature. Some researchers, like Butterworth (2010), define it based on persistent struggles with numbers and operations, while others, such as Price and Ansari (2013), conceptualize it as a deficit in neurological skills. A more comprehensive view from Kaufman et al. (2011) describes MLD as a combination of deficits across behavioral, cognitive, neuropsychological, and neural domains. Despite these different perspectives, the common characteristic is a significant and enduring difficulty with understanding, recalling, and applying mathematical concepts or correctly using operational rules. To create a unified conceptualization for this research, the current study synthesizes these viewpoints through the lens of the diagnostic criteria from DSM-5 and ICD-10.

Diagnostic Criteria for Mathematics Learning Disabilities

Globally, the diagnostic approach to MLD is guided by the standards set forth by the American Psychiatric Association (APA) in the DSM-5 and the World Health Organization (WHO) in the ICD-10 (APA, 2013; Kişiler et al., 2021; Pham & Riviere, 2015; Tannock, 2013). These frameworks outline three core cognitive dimensions of the disability: number sense, calculation skills, and mathematical reasoning. As defined by Kaufmann & Von Aster (2012), number sense involves challenges in understanding fundamental mathematical ideas, magnitudes, and symbols. Concurrently, deficits in computational skills refer to an inability to execute basic arithmetic or a high frequency of errors. The third dimension, mathematical reasoning, pertains to difficulties with the application of mathematical knowledge to problem-solving.

A diagnosis of MLD must consider its emergence from an interplay of cognitive and academic factors, necessitating an evaluation of number sense, arithmetic, and problem-solving. Both number sense (which includes skills like counting and number comparison) and basic arithmetic (addition, subtraction, multiplication, and division) are widely viewed as critical markers for identifying the disability (Mutlu, 2019; 2024). These particular skills are foundational, forming the bedrock for higher-level mathematics (Gersten & Chard, 1999). Consequently, students who struggle with fundamental numerical skills often fail to achieve arithmetic fluency—the ability to perform calculations with speed and accuracy (Molise & Luneta, 2024). Since individuals with MLD often show difficulties throughout the problem-solving process (Monei & Pedro, 2017), it is essential that problem-solving capabilities are also a component of the diagnostic examination.

Assessment of Mathematics Learning Disabilities

The assessment of MLD is complicated by its frequent co-occurrence with other learning disabilities, which necessitates the use of specialized diagnostic tools. While traditional intelligence and academic achievement tests are often employed for this purpose (Geary, 2004; Eteng-Uket, 2023; Geary & Hoard, 2005), their effectiveness can be inadequate. These tests are prone to being skewed by socioeconomic factors and often neglect the emotional and environmental dimensions of the disability (Gersten

et al., 2005; Mazzocco & Myers, 2003; Yoong et al., 2023), thereby highlighting a clear need for more sensitive instruments. In response to these limitations, a wide array of alternative tools has emerged in the literature. Several instruments focus on screening, such as the Number Sets Test for identifying risk in young children (Geary et al., 2009), a computer-based tool from Butterworth (2003) for students aged 6-14, and Loughborough University's "dyscalculium" screener (Eteng-Uket, 2023). More specific diagnostic instruments include the Neuropsychological Test Battery (NUCALC), which evaluates number processing and calculation (von Aster & Shalev, 2007), and the TEDIMATH test for diagnosing arithmetic disorders (Grégoire et al., 2004, as cited in Eteng-Uket, 2023). Other researchers have designed computer-based tests for basic mathematical and psychological processes (Cangoz et al., 2013; Karagiannakis & Baccaglini-Frank, 2014). Despite this progress, many of these tools, as well as methods like discrepancy analysis and behavioral observation, are criticized for lacking standardization, having psychometric weaknesses, and maintaining too narrow a focus (Cangoz et al., 2013; Eteng-Uket, 2023; Mutlu & Akgün, 2017). This critique is reinforced by the findings of Mutlu and Akgün (2017), who determined that all common diagnostic approaches have significant limitations when used alone. Given the profound impact of MLD on a child's development and the difficulty of early diagnosis, creating a reliable, effective, and teacher-friendly screening tool is a vital step toward supporting students and educators.

Beyond these general assessment challenges, a specific problem exists in Türkiye. While numerous MLD assessment tools are available internationally, there is a scarcity of scales that are culturally and linguistically validated for the Turkish context. Many existing scales are direct translations, which raises concerns about their cross-cultural validity—specifically, whether they measure the same cognitive constructs in individuals from different cultures (Cangoz et al., 2013; Geary, 2004; Mutlu & Akgün, 2017; Filiz, 2021). This concern is substantiated by research indicating that the cognitive manifestations of learning difficulties can differ across language groups (e.g., English vs. Turkish learners) (Altunkaya, 2017; Güngör & Yeşilyurt, 2023).

While there have been recent efforts to develop assessment scales in Türkiye, the demand for such tools has not been satisfied. This gap is attributable to the demanding nature of scale validation and the difficulty of creating a comprehensive assessment that reflects the complex, multidimensional aspects of MLD. In response, this study's primary purpose is to develop the Mathematics Learning Disability Screening Scale (MLDSS), a valid, reliable, and culturally-sensitive tool for the Turkish context. The broader goals of this work are to refine the definition of the MLD population in Türkiye, highlight current inadequacies, and provide a critical resource for teachers and researchers that aids in the early identification and support of at-risk students. The entire study is structured around answering two fundamental research questions about the proposed scale:

1. Does the Mathematics Learning Disability Screening Scale (MLDSS) demonstrate validity in measuring the symptoms of students with mathematics learning disabilities?
2. Does the Mathematics Learning Disability Screening Scale (MLDSS) demonstrate reliability in measuring the symptoms of students with mathematics learning disabilities?

METHOD

Research Design

This study evaluates the psychometric properties of the Mathematics Learning Disability Screening Scale (MLDSS), which was developed for elementary and secondary school students. Accordingly, this research is a scale development study conducted within the framework of a survey research design. Survey research designs involve collecting data on characteristics such as attitudes and behaviors from a large sample to describe the current state of a population (Fraenkel et al., 2015; McMillan & Schumacher, 2014). This study involved administering the scale to elementary and secondary school students, followed by an examination of its reliability, validity, factor structure, and other psychometric parameters. Specifically, an Exploratory Factor Analysis (EFA) was first conducted to identify the underlying structure of the scale. Subsequently, a Confirmatory Factor Analysis (CFA) was performed to validate this structure. The results of these analyses are reported sequentially.

Participants

To achieve a sample that was representative of the target population, this study employed cluster sampling, a probability-based technique (Fraenkel et al., 2015). The procedure for this method involved identifying appropriate sampling units (clusters) and subsequently selecting participants at random from within those units (Karagöz, 2023). The data collection effort spanned all seven of Türkiye's geographical regions, with two provinces chosen from each, totaling fourteen. The process began by identifying primary and secondary schools with a high incidence of students with learning disabilities. From these schools, teachers who instructed a student with a formal diagnosis of a learning disability were invited to take part in the study. Teachers who volunteered were then tasked with completing the rating scale for a single student under their supervision.

In total, 644 teachers from 120 schools participated, completing the scale for 644 students who had been diagnosed with a learning disability. The ages of the students ranged from 9 to 12. To ensure high representativeness, data were collected from an equal or near-equal number of students from each grade level. According to the literature on factor analysis, sample sizes of 100 are considered weak, 200 fair, 300 good, 500 very good, and 1,000 excellent (Field, 2013). Given the number of participants in this study (N=644), the sample size can be considered very good for conducting a factor analysis. Detailed demographic information about the participants is presented below.

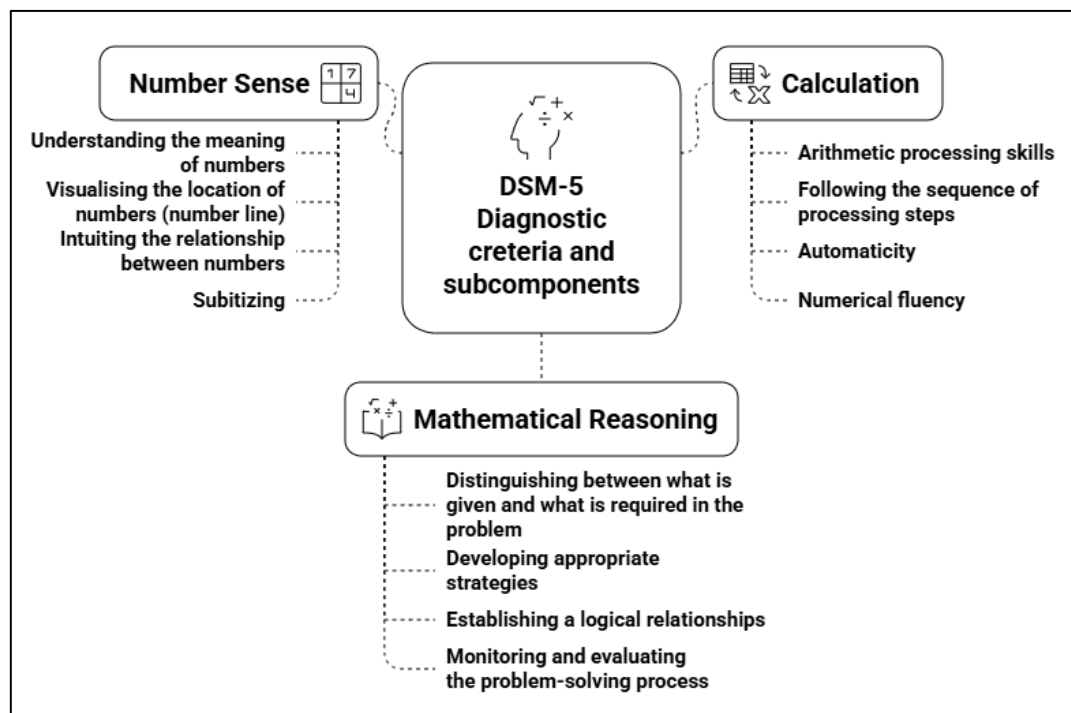
Table 1. Demographic characteristics of the participants

Variable	Category	EFA Sample		CFA Sample	
		Frequency	Percentage	Frequency	Percentage
Gender	Female	206	%60	175	58.3
	Male	138	%40	125	%41.7
	Total	344	%100	300	%100
Grade	3	90	%26.1	78	%26
	4	82	%23.9	71	%23.7
	5	84	%24.4	82	%27.3
	6	88	%25.6	69	%23
	Total	344	%100	300	%100

Procedure

This section outlines the steps followed in the development of the scale. The process adhered to the standard procedures for scale development as reported in the literature (DeVellis & Thorpe, 2022; Simms, 2008). The examination of the psychometric properties of the MLDSS developed in this study followed these established standards.

First, the construct of MLD and its components were defined. To this end, the MLD literature was analyzed in detail, along with the diagnostic criteria for learning disabilities (LDs) and MLD from DSM-5 and ICD-10. A three-dimensional heterogeneous structure reflecting the nature of MLD was identified. This structure was classified under three categories (see Fig 1): Number Sense, Calculation, and Mathematical Reasoning.

**Figure 1. DSM-5 Diagnostic criteria and subcomponents**

The second phase of the procedure involved creating the draft scale, a process guided by the three sequential stages of item writing, expert review, and a pilot study (Tavşancıl, 2014). Initially, to build a comprehensive profile of MLD, an initial pool of 62 items was generated through a thorough literature search of electronic databases, including TÜBİTAK ULAKBİM (Turkish Academic Network and Information Center), Web of Science, Scopus, Google Scholar, and Science Direct (APA, 2013; Geary et al., 2009; Hellstrand et al., 2024; etc.). This pool was then refined to 51 items after redundant items were eliminated. To establish content validity, the 51-item draft was reviewed by an expert panel consisting of a psychologist, a special education specialist, a mathematics teacher, and two MLD researchers. The panel evaluated the scale's scope and sub-dimensions, as well as the clarity, wording, and psychometric appropriateness of each item. Although the experts' overall feedback was positive, confirming the scale's theoretical soundness and suitability, they recommended several changes. Consequently, items with similar wording (M 23, 33, 48) and those that did not align with DSM-5 criteria and the literature (M 1, 2, 3, 5, 8, 11, 12, 13, 15, 16, 20, 24, 25) were removed. This review process resulted in a final 35-item version, for which a five-point Likert-type rating scale ("never" to "always") was adopted for responses. This final version was then administered in a pilot study to 20 teachers (10 primary school,

10 mathematics) outside of the main study group, and the successful pilot confirmed the items were clear and understandable, preparing the scale for full administration.

Third, before completing the MLDSS, all participating teachers were directed to attend an online training session prepared by the researchers. This two-hour training was conducted via videoconference on Microsoft Teams and was recorded and shared with the teachers via Google Drive. The training aimed to inform teachers about MLD and to encourage their active participation in the data collection process. The content covered topics such as the definition and diagnosis of MLD, a review of the MLDSS items, instructions on how to administer the scale, and teaching methods and interventions for students with MLD. Additionally, a pre-assessment session was held to address any challenges faced by the teachers (e.g., accessing or filling out the Google Form). The scale was administered online using Google Forms, and teachers were given a two-month period to observe and evaluate the performance of their students.

The fourth and final procedural step was to conduct validity and reliability analyses of the MLDSS, for which data were gathered from two separate samples: 344 teachers for an Exploratory Factor Analysis (EFA) and 300 for a Confirmatory Factor Analysis (CFA). EFA, performed in SPSS 25, was utilized to identify the underlying factor structure, analyze how items grouped into sub-dimensions, and examine their loadings and inter-factor correlations (DeVellis, 2003). Following this, CFA, conducted in AMOS 23, was used to confirm whether the proposed factor structure was supported by the data from the second sample, a key step in establishing construct validity (Brown, 2015; Kline, 2016). To assess reliability, the MLDSS was also re-administered to 20 teachers after two weeks to establish test-retest stability via the Pearson Correlation Coefficient, while internal consistency was measured using Cronbach's Alpha. The study concluded with a comprehensive interpretation and discussion of all findings related to the scale's psychometric properties.

FINDINGS

Findings Related to Validity Analysis

This section presents the findings related to the research question: "Does the MLDSS validly measure the symptoms of students with mathematical learning disabilities?" To examine the construct validity of the MLDSS, Exploratory Factor Analysis (EFA) and Confirmatory Factor Analysis (CFA) were employed.

Findings Related to EFA

To test the hypothesized three-dimensional structure of the scale, which was based on DSM-5 criteria, an Exploratory Factor Analysis (EFA) was conducted. Before beginning the main analysis, the suitability of the data for factor analysis was confirmed. The Kaiser-Meyer-Olkin (KMO) measure yielded a value of .931, indicating marvelous sample adequacy (Kaiser, 1974). Furthermore, Bartlett's Test of Sphericity was significant ($\chi^2 = 3011.274$; $df = 120$; $p < .001$), which confirmed that the correlations among the items were strong enough for analysis (Bartlett, 1954). Because the underlying factors were assumed to be correlated, the Direct Oblimin rotation method was selected, as is common in educational research.

The analysis then proceeded with an iterative item reduction process. The EFA was first conducted on 35 items, which initially produced a three-factor structure explaining 60.115% of the total variance. In the first refinement stage, items were removed for statistical reasons: those with low factor loadings below 0.40 (M 14, 21, 22, 29, 30, 31) and those with high cross-loadings on multiple factors (M 41, 44, 46, 49, 51). After re-running the analysis, a second refinement was performed, where items were reviewed for their theoretical contribution. Items that did not seem to measure the intended construct (M 7, 18, 34, 37, 42, 45, 50) and one additional item with a low factor loading (M 9) were also removed. This iterative process resulted in a final 16-item scale, for which a final EFA revealed a three-factor structure that explained 68.495% of the total variance. The detailed results for this final structure are presented in Table 2.

Table 2. Factor characteristics

Factor	Eigenvalue	Variance (%)	Cumulative variance (%)
Number sense	8.057	50.353	50.353
Calculation	1.882	11.763	62.117
Mathematical reasoning	1.021	6.379	68.495

An examination of Table 2 shows that, according to the EFA results, all three sub-dimensions of the MLDSS have eigenvalues greater than 1.0. The first sub-dimension has a particularly high eigenvalue of 8.057, explaining 50.353% of the variance. The second sub-dimension has an eigenvalue of 1.882 and accounts for an additional 11.763% of the variance. Finally, the third sub-dimension has an eigenvalue of 1.021, explaining a further 6.379% of the variance. To support the analysis of the scale's sub-dimensions, the scree plot is presented in Figure 2.

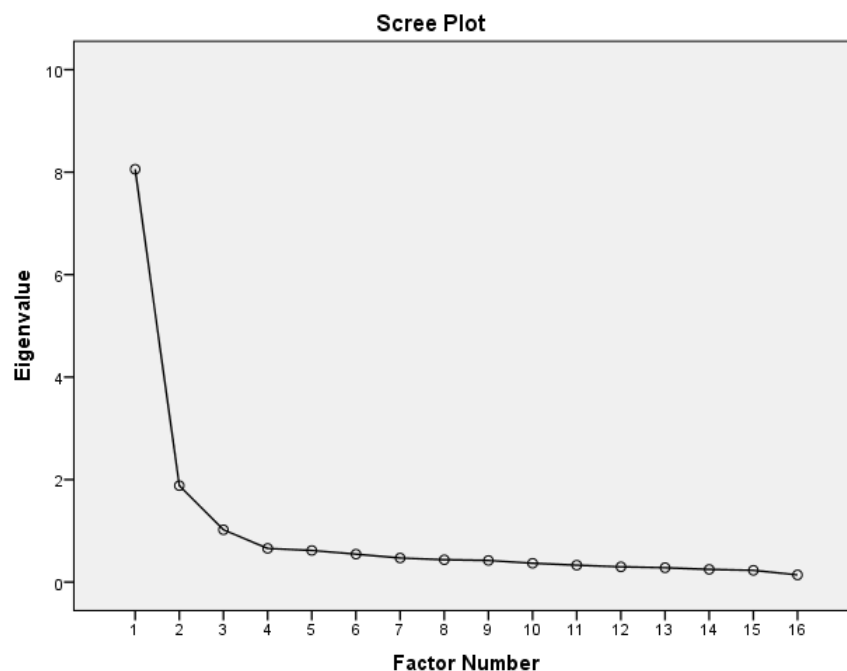


Figure 2. Scree plot

To determine the number of factors to retain, an inspection of the scree plot was conducted (see Fig 1). The graph shows a clear "elbow" forming between the third and fourth factors, after which the eigenvalues decrease sharply and begin to level off. This pattern indicates that the initial three factors explain a significant amount of the variance, while any subsequent factors contribute very little explanatory power. Based on this visual analysis, an optimal three-factor solution was concluded for the scale. To further substantiate this decision, the proposed structure was reviewed by a panel of experts, including two internationally recognized faculty members with expertise in MLD and one specialist in scale development. After evaluating the factors and their corresponding items, the experts endorsed the three-factor structure as theoretically sound. Therefore, the analysis proceeded with this confirmed three-factor model, and the results of the subsequent rotated factor analysis are detailed in Table 3.

Table 3. Results of the rotated factor analysis

Factor Name		Item	Communality (Extraction)	Factor Loadings		
				1	2	3
Number Sense	M4	Inverts or changes the order of numbers when reading and writing (for example, writes or reads 69 as 96 or 342 as 234)	0,69	0,90		
	M38	Writes numbers to the wrong digit in mathematical operations	0,72	0,86		
	M35	Difficulties in determining the place value (especially numbers with zero tens digit in three-digit numbers (for example, 305))	0,69	0,83		
	M40	Difficulties in establishing a one-to-one relationship between number symbols and objects when counting	0,70	0,81		
	M39	Has difficulty rounding numbers to tens or hundreds (especially numbers with 5 in ones and tens digits)	0,70	0,73		
	M43	Difficulties in rhythmic counting backwards starting from any number	0,53	0,68		
	M47	Difficulties in comparing numbers (e.g. less, more)	0,47	0,49		
	M32	Difficulty in showing the position of any number on the number line	0,54	0,45		
Calculation	M26	Has difficulty in calculating the total amount when shopping	0,82		-0,93	
	M27	Has difficulty in calculating change when paying for shopping	0,83		-0,88	
	M28	Has difficulty in reading an analogue clock	0,54		-0,53	
Mathematical Reasoning	M6	Has difficulty in using different problem solving strategies when solving problems	0,43			0,70
	M4	Difficulties in interpreting mathematical problems	0,49			0,70
	M17	Difficulties in mathematical thinking	0,57			0,74
	M10	Difficulties in interpreting numerical data	0,48			0,59
	M19	Difficulties in concretising and visualising abstract mathematical concepts (modelling)	0,49			0,66

As shown in Table 3, the communalities for each item ranged from .45 to .93. The scale consists of three sub-dimensions: "Number Sense," comprising 8 items; "Calculation," comprising 3 items; and "Mathematical Reasoning," comprising 5 items. The naming of these dimensions was based on the DSM-5 classification system and expert opinions. The factor loadings for the items in the "Number Sense" sub-dimension ranged from .45 to .90. For the "Calculation" sub-dimension, the factor loadings ranged from -.93 to -.55. For the "Mathematical Reasoning" sub-dimension, the factor loadings ranged from .62 to .74. Furthermore, an inter-factor correlation analysis was conducted to determine the relationships among the scale's factors (see Table 4).

Table 4. Inter-factor correlation analysis

Factors	1	2	3
Number Sense (1)	1.00		
Calculation (2)	-.681	1.00	
Mathematical Reasoning (3)	.508	-.449	1.00

**p<.01

As detailed in Table 4, the analysis of the scale's factors showed moderate and significant correlations. Specifically, Number Sense was negatively correlated with Calculation ($r = -.681$, $p < .001$) but positively correlated with Mathematical Reasoning ($r = .508$, $p < .001$). A further negative correlation was identified between Calculation and Mathematical Reasoning ($r = -.449$, $p < .001$). The strength of these correlations sufficiently demonstrates that the sub-dimensions measure related but distinct constructs without issues of multicollinearity.

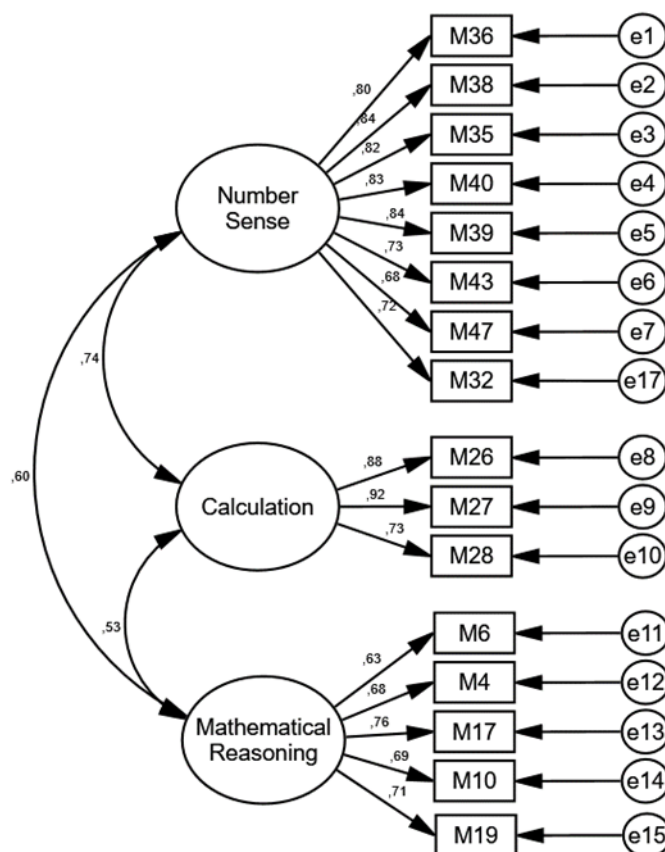
Findings Related to CFA

To evaluate the model fit from the Confirmatory Factor Analysis (CFA), this study employed a range of goodness-of-fit indices, an approach necessitated by the known limitations of the chi-square (χ^2) test with large samples. Therefore, the model was evaluated against established criteria using both absolute indices (χ^2/df , GFI, AGFI, SRMR, RMSEA) and comparative indices (CFI, IFI). The thresholds for acceptable fit included a χ^2/df ratio below 5.0 (Wheaton et al., 1977), GFI/AGFI values above .85, SRMR/RMSEA values below .08 (Hooper et al., 2008), and CFI/IFI values above .90 (Bentler, 1995; Bollen, 1989). Table 5 provides the complete findings for these goodness-of-fit indicators.

Table 5. CFA Goodness-of-Fit Indices for the MLDSS.

Fit Index	Value	Acceptable Fit	Excellent Fit	Status
NFI	.92	$\geq .90$	$\geq .95$	Acceptable
NNFI	.92	$\geq .90$	$\geq .95$	Acceptable
IFI	.95	$\geq .90$	$\geq .95$	Perfect
CFI	.95	$\geq .90$	$\geq .95$	Perfect
RFI	.90	$\geq .90$	$\geq .95$	Acceptable
AGFI	.87	$\geq .85$	$\geq .90$	Acceptable
GFI	.90	$\geq .85$	$\geq .90$	Perfect
RMR	.044	$\leq .08$	$\leq .05$	Perfect
SRMR	.051	$\leq .08$	$\leq .05$	Perfect
REMSEA	.069	$\leq .08$	$\leq .05$	Acceptable
χ^2/df	2.41	2.5-3	0-2.5	Perfect

An analysis of the goodness-of-fit indices presented in Table 5 indicates that the values for IFI, CFI, GFI, RMR, SRMR, and χ^2/df demonstrate an excellent model fit. The values for NFI, NNFI, RFI, AGFI, and RMSEA indicate an acceptable model fit. The R^2 values, which represent the proportion of variance in each item explained by its factor, ranged from .11 to .59. To further examine the model's structural fit, the path diagram illustrating item-total correlations and the corresponding t-values was inspected. The path diagram for the first-order CFA is presented in Figure 3.



CMIN/df:2,419; AGFI:,.876; GFI:,.908; NFI:,.921; CFI:,.951; IFI:,.952; TLI:,.942; RMSEA:,.069

Figure 3. Path diagram

The path diagram in Figure 3 illustrates that all items loaded strongly onto their respective factors, with standardized coefficients ranging from .63 to .92. Because all factor loadings also had significant t-values ($p < .01$), the findings indicate that the scale's items are effective and distinct measures of their underlying constructs.

Findings Related to Reliability Analysis

The reliability of the MLDSS was confirmed through two separate analyses. First, the scale's internal consistency was found to be excellent, as measured by Cronbach's Alpha. The alpha for the total scale was .93 in both the EFA and CFA samples, with high coefficients also observed for the sub-dimensions in each sample (EFA sample: .89, .89, and .84; CFA sample: .92, .87, and .82 for Number Sense, Calculation, and Mathematical Reasoning, respectively). Second, the scale's stability was affirmed through a test-retest analysis with 20 teachers across a 15-day period. This yielded strong Pearson correlation coefficients for all sub-dimensions—Number Sense (.90), Calculation (.92), and Mathematical Reasoning (.88)—as well as for the total scale score (.90), indicating its consistency over time.

DISCUSSION AND CONCLUSION

This study was designed to develop the Mathematics Learning Disability Screening Scale (MLDSS), a valid and reliable instrument based on DSM-5 diagnostic criteria for identifying elementary school students at risk for mathematics learning disabilities (MLD). The findings demonstrate that the developed scale possesses strong psychometric properties, showing both high validity and reliability. These results align closely with current international and national literature, indicating that the MLDSS offers a significant contribution to the field.

The construct validity analysis confirmed a three-dimensional theoretical structure for MLD, comprising number sense, calculation, and mathematical reasoning. This multidimensional framework corroborates the scientific consensus that MLD arises from fundamental deficits in interrelated yet distinct skill areas, rather than from a single, monolithic problem. This structure is conceptually consistent with findings from other international instruments. For example, the "Dyscalculia Test" developed in Nigeria by Eteng-Uket (2023) also focuses on three core domains: number sense, arithmetic operations, and working memory. In contrast, a mathematics achievement test developed for elementary school students in Hong Kong approached MLD assessment through more curriculum-based domains such as mathematical fluency, arithmetic, and problem-solving (Kwan, 2020). The MLDSS

is uniquely positioned within this landscape as a tool that targets not only the academic manifestations of MLD but also its cognitive roots, particularly by treating a core deficit like number sense as a distinct dimension.

The psychometric properties of the MLDSS further attest to its robustness. The scale's three-factor structure explains a total variance of 68.5%, a notably high value that indicates a strong model fit (Zahra et al., 2014). Furthermore, the high internal consistency (Cronbach's $\alpha = .93$) and test-retest reliability (.90) underscore the instrument's stability and precision. These reliability coefficients are comparable to, and in some cases exceed, those of other well-regarded scales in both international (Geiger & Brewster, 2018; Eteng-Uket, 2023) and national literature (Okur, 2019; Deniz, 2022).

A central finding from the MLDSS development is the pivotal role of the "number sense" factor in explaining MLD. This corroborates a large body of research suggesting that a core deficit in understanding and processing numerical quantities lies at the heart of MLD. For instance, using the Number Sets Test, Geary et al. (2009) demonstrated that number sense performance was one of the strongest predictors of MLD, independent of intelligence and working memory. Similarly, Olkun et al. (2016), in their work on an MLD screening tool in Türkiye, highlighted that deficits in basic number processing skills are a elementary reason for low mathematics achievement. The ability of the MLDSS to successfully measure number sense as a discrete factor suggests it has strong potential to detect this critical cognitive deficit.

These findings also align with research demonstrating the feasibility of early and reliable identification of children at risk for MLD. Mazzocco and Thompson (2005) found that kindergarteners' performance on basic numeracy tasks could predict a later MLD diagnosis with high accuracy. This reinforces the call from researchers like Brendefur et al. (2018) for more comprehensive assessment tools that evaluate multiple underlying domains, rather than focusing solely on number concepts. The importance of this study is further highlighted by the global trend towards culturally and contextually sensitive assessment. The development of a local test in Hong Kong (Kwan, 2020) mirrors the need for the MLDSS in Türkiye, pointing to a universal recognition that such instruments must be localized. The MLDSS serves this purpose by measuring performance in specific domains directly linked to MLD (number sense, calculation, and reasoning), thereby providing a valuable profile that can inform subsequent diagnostic processes.

RECOMMENDATIONS

The MLDSS presents significant implications for both educational practice and research.

For Practitioners: The scale can serve as a critical tool for teachers and school psychologists in the early identification of students at risk for MLD. By providing concrete data on a student's specific area of weakness (e.g., number sense versus calculation), the MLDSS enables the design of targeted and effective educational programs. This approach aligns with the philosophy of tools like the PMA (Brendefur et al., 2018), which advocate for interventions tailored to a student's specific needs rather than a "one-size-fits-all" approach.

For Researchers: The MLDSS can be utilized as a foundational data collection tool for studies investigating the cognitive profiles, prevalence, and longitudinal progression of MLD in Türkiye. Future longitudinal studies examining the predictive power of the MLDSS, similar to the work of Geary et al. (2009), would further solidify its diagnostic value. Moreover, conducting comparative analyses with other local instruments (e.g., Olkun et al., 2016) and establishing norms with larger, geographically diverse samples would enhance the scale's generalizability and contribute significantly to the body of knowledge in this field.

Declaration of Conflicting Interests

The authors declared no potential conflicts of interest with respect to the research, authorship or publication of this article.

Funding

The authors received no financial support for the research, author-ship, and publication of this article.

Statements of publication ethics

We hereby declare that the study has not unethical issues and that research and publication ethics have been observed carefully.

Researchers' contribution rate

TF and AÇ conceived of the presented idea and jointly developed the items for the scale. TF designed the methodology and performed the statistical analyses. AÇ provided the theoretical framework and supervised the interpretation of the findings. TF prepared the Method and Findings sections, while AÇ wrote the Introduction and Discussion. Both authors managed the data collection and ethics approval processes. All authors discussed the results and contributed to the final manuscript.

Ethics Committee Approval Information

This study was approved by the Bayburt University Social and Human Sciences Research Ethics Committee (Decision No: 28, Date: 07/01/2025).

REFERENCES

- Akın, A., & Sezer, S. (2010). Diskalkuli: Matematik öğrenme bozukluğu. *Bilim ve Aklın Aydınlığında Eğitim*, 126, 41-48.
- Altunkaya, H. (2017). Yabancı dil olarak Türkçe öğrenenlerin dinleme ve okuma kaygıları. *Education Sciences*, 12(3), 107-121. <https://doi.org/10.12739/NWSA.2017.12.3.1C0672>
- Asfuroğlu, B., & Fidan, S. (2016). Özgül öğrenme güçlüğü/spesifik learning disorders. *Osmangazi Tıp Dergisi*, 38(1), 49-54. <http://dx.doi.org/10.20515/otd.17402>
- Bartlett, M. S. (1954). A note on the multiplying factors for various χ^2 approximations. *Journal of the Royal Statistical Society Series B (Methodological)*, 16(2), 296-298.
- Bland, J. M., & Altman, D. G. (1997). Statistics notes: Cronbach's alpha. *Bmj*, 314(7080), 572. <https://doi.org/10.1136/bmj.314.7080.572>
- Brendefur, J. L., Johnson, E. S., Thiede, K. W., Strother, S., & Severson, H. H. (2018). Developing a multi-dimensional early elementary mathematics screener and diagnostic tool: the primary mathematics assessment. *Early Childhood Education Journal*, 46(2), 153-157. <https://doi.org/10.1007/s10643-017-0854-x>
- Brown, T. A. (2015). *Confirmatory factor analysis for applied research*. Guilford publications.
- Butterworth, B. (2003). Dyscalculia Screener. 1-74. <https://doi.org/10.1037/t05204-000>
- Butterworth, B. (2010). Foundational numerical capacities and the origins of dyscalculia. *Trends in cognitive sciences*, 14(12), 534-541.
- Cangoz, B., Altun, A., Olkun, S., & Kacar, F. (2013). Computer based screening dyscalculia: cognitive and neuropsychological correlates. *Turkish Online Journal of Educational Technology-TOJET*, 12(3), 33-38.
- Cirino, P. T., Fuchs, L. S., Elias, J. T., Powell, S. R., & Schumacher, R. F. (2015). Cognitive and mathematical profiles for different forms of learning difficulties. *Journal of Learning Disabilities*, 48(2), 156-175. <https://doi.org/10.1177/0022219413494239>
- Çokluk, Ö., Şekercioğlu, G., ve, & Büyüköztürk, Ş. (2018). *Sosyal bilimler için çok değişkenli istatistik: SPSS ve LISREL uygulamaları* (5th ed.). Pegem Akademi.
- Deniz, S. (2022). Development of Specific Learning Disability Screening Scale. *International Journal of Progressive Education*, 18(6), 166-179.. <https://doi.org/10.29329/iipe.2022.477.11>
- Deruaz, M., Dias, T., Gardes, M. L., Gregorio, F., Ouvrier-Buffet, C., Peteers, F., & Robotti, E. (2020). Exploring MLD in mathematics education: Ten years of research. *The Journal of Mathematical Behavior*, 60, 100807. <https://doi.org/10.1016/j.jmathb.2020.100807>
- DeVellis, R. F. (2003). *Scale development theory and applications* (2nd ed.). SAGE Publications International Educational and Professional Publisher.
- DeVellis, R. F., & Thorpe, C. T. (2022). *Scale development: Theory and applications* (5th ed.), Sage.
- Eteng-Uket, S. (2023). The development, validation, and standardization of a new tool: The Dyscalculia Test. *Numeracy*, 16(2), 1-28. <https://doi.org/10.5038/1936-4660.16.2.1417>
- Field, A. (2013). *Discovering statistics using IBM SPSS statistics* (4th ed.). Sage.
- Filiz, T. (2021). Matematik öğrenme güçlüğü yaşayan öğrencilere yönelik öğretimsel müdahalelerin öğrencilerin akademik başarılarına etkisinin incelenmesi. *Ankara Üniversitesi Eğitim Bilimleri Fakültesi Özel Eğitim Dergisi*, 22(4), 1025-1055. <https://doi.org/10.21565/oelegitimdergisi.713496>
- Filiz, T., & Güneş, G. (2023). Teaching Multiplication Based on Realistic Mathematics Education to Students at Risk of Mathematics Learning Difficulties. *Journal of Family Counseling and Education*, 8(2), 193-211. <https://doi.org/10.32568/jfce.1372345>
- Fraenkel, J.R., Wallen, N.E., & Hyun, H.H. (2015). *How to design and evaluate research in education* (9th ed.). Mc Graw Hill Education.
- Geary, D. C. (2004). Mathematics and learning disabilities. *Journal of Learning Disabilities*, 37(1), 4-15. <https://doi.org/10.1177/00222194040370010201>
- Geary, D. C. (2013). Early foundations for mathematics learning and their relations to learning disabilities. *Current Directions in Psychological Science*, 22(1), 23-27. <https://doi.org/10.1177%2F0963721412469398>
- Geary, D. C., & Hoard, M. K. (2005). Learning disabilities in arithmetic and mathematics: Theoretical and empirical perspectives. In *The handbook of mathematical cognition* (pp. 253-267). Psychology Press.
- Geary, D. C., Bailey, D. H., & Hoard, M. K. (2009). Predicting mathematical achievement and mathematical learning disability with a simple screening tool: The number sets test. *Journal of psychoeducational assessment*, 27(3), 265-279. <https://doi.org/10.1177/0734282908330592>
- Geiger, E. F., & Brewster, M. E. (2018). Development and evaluation of the individuals with learning disabilities and/or difficulties perceived discrimination scale. *The Counseling Psychologist*, 46(6), 708-737. <https://doi.org/10.1177/0011000018794919>
- Gersten R., N. C. Jordan, & J. R. Flojo. (2005). Early identification and interventions for students with mathematics difficulties. *Journal of Learning Disabilities*, 38, 293-304. <https://doi.org/10.1177/00222194050380040301>
- Gersten, R., & Chard, D. (1999). Number sense: Rethinking arithmetic instruction for students with mathematical disabilities. *The Journal of Special Education*, 33(1), 18-28.
- Güngör, H., & Yeşilyurt, Ş. (2023). İki Dilli Okuma Metinlerinin Okuduğunu Anlama Başarısına Etkisi. *Anadolu Dil ve Eğitim Dergisi*, 1(1), 1-11. <https://10.5281/zenodo.7772503>
- Hellstrand, H., Holopainen, S., Korhonen, J., Räsänen, P., Hakkarainen, A., Laakso, M.-J., Laine, A., & Aunio, P. (2024). Arithmetic fluency and number processing skills in identifying students with mathematical learning disabilities. *Research in Developmental Disabilities*, 151, 104795. <https://doi.org/10.1016/j.ridd.2024.104795>
- Hooper, D., Coughlan, J., & Mullen, M. (2008). Evaluating model fit: a synthesis of the structural equation modelling literature. In *7th European Conference on research Methodology for Business and Management Studies* 195-200. <https://doi.org/10.21427/D79B73>
- Kaiser, H. F. (1974). An index of factorial simplicity. *Psychometrika*, 39(1), 31-36.
- Karagiannakis, G., & Baccaglini-Frank, A. (2014). The DeDiMa battery: A tool for identifying students' mathematical learning profiles. *Health Psychology Report*, 2(4), 291-297. <https://doi.org/10.5114/hpr.2014.46329>
- Kaufmann, L., & von Aster, M. (2012). The diagnosis and management of dyscalculia. *Deutsches Ärzteblatt International*, 109(45), 767. <https://doi.org/10.3238%2Faztebl.2012.0767>
- Kaufmann, L., Wood, G., Rubinsten, O., & Henik, A. (2011). Meta-analyses of developmental fMRI studies investigating typical and atypical trajectories of number processing and calculation. *Developmental neuropsychology*, 36(6), 763-787. <https://doi.org/10.1080/87565641.2010.549884>

- Kiřler, C., Schwenk, C., & Kuhn, J. T. (2021). Two dyscalculia subtypes with similar, low comorbidity profiles: A mixture model analysis. *Frontiers in Psychology*, 12, 589506. <https://doi.org/10.3389/fpsyg.2021.589506>
- Kline, R. B. (2016). *Structural equation modeling* (4th ed.). Guilford.
- Kwan, K. T. (2020). Developing a math achievement test for identifying primary students at-risk for dyscalculia in Hong Kong [Master's thesis, The Education University of Hong Kong].
- Mazzocco, M. M. M., & Myers, G. F. (2003). Complexities in identifying and defining mathematics learning disability in the primary school-age years. *Annals of Dyslexia*, 53(1), 218–253. <https://doi.org/10.1007/s11881-003-0011-7>
- Mazzocco, M. M., & Thompson, R. E. (2005). Kindergarten predictors of math learning disability. *Learning Disabilities Research & Practice*, 20(3), 142–155. <https://doi.org/10.1111/j.1540-5826.2005.00129.x>
- McMillan, J., & Schumacher, S. (2014). *Research in education: Evidence-based inquiry* (7th ed.). Pearson Education Limited.
- Molise, D. C., & Luneta, K. (2024). An overview of the causes of dyscalculia and its impact on learners' arithmetic ability. *The Independent Journal of Teaching and Learning*, 19(1), 124–144. <http://dx.doi.org/10.17159/ijtl.v19i1.18857>
- Monei, T., & Pedro, A. (2017). A systematic review of interventions for children presenting with dyscalculia in primary schools. *Educational Psychology in Practice*, 33(3), 277–293. <https://doi.org/10.1080/02667363.2017.1289076>
- Mutlu, Y. (2019). Math anxiety in students with and without math learning difficulties. *International Electronic Journal of Elementary Education*, 11(5), 471–475. <https://doi.org/10.26822/iejee.2019553343>
- Mutlu, Y. (2024). Effects of dyscalculia on personal, social, academic, professional and daily life: A case study. *International Electronic Journal of Elementary Education*, 17(1), 89–101. <http://dx.doi.org/10.26822/iejee.2024.365>
- Mutlu, Y., & Akgün, L. (2017). Matematik öğrenme güçlüğü tanılamada yeni bir model önerisi: Çoklu süzgeç modeli. *İlköğretim online*, 16(3), 1153–1173. <http://dx.doi.org/10.17051/ilkonline.2017.330247>
- Mutlu, Y., Çam, Z., & Çalışkan, E. F. (2021). *Diskalkulik çocukların yaşadıkları sorunlar ve çözüm önerileri*. Vize yayıncılık.
- Okur, M. (2019). Öğrenme güçlüğü erken belirtileri tarama ölçeğinin psikometrik niteliklerinin belirlenmesi. [Yüksek Lisans Tezi, Anadolu Üniversitesi].
- Olkun, S. (2015). Matematik öğrenme güçlükleri. S. Sunay Yıldırım-Doğru (Ed.), *Öğrenme güçlükleri içinde* (s. 211–226). Eğiten Kitap.
- Olkun, S., Altun, A., Gocer Sahin, S., & Kaya, G. (2016). Psychometric properties of a screening tool for elementary school student's math learning disorder risk. *Online Submission*, 15(12), 48–66.
- Pham, A. V., & Riviere, A. (2015). Specific learning disorders and ADHD: Current issues in diagnosis across clinical and educational settings. *Current Psychiatry Reports*, 17, 1–7. <https://doi.org/10.1007/s11920-015-0584-y>
- Price, G. R., & Ansari, D. (2013). Dyscalculia: Characteristics, causes, and treatments. *Numerical Cognition: A Volume in Honor of Bert Reynvoet*, 145–172. <http://dx.doi.org/10.5038/1936-4660.6.1.2>
- Simms, L. J. (2008). Classical and modern methods of psychological scale construction. *Social and Personality Psychology Compass*, 2(1), 414–433. <https://doi.org/10.1111/j.1751-9004.2007.00044.x>
- Tabachnick, B. G., & Fidell, L. S. (2013). *Using multivariate statistics* (6th ed.). Pearson.
- Tannock, R. (2013). Rethinking ADHD and LD in DSM-5: Proposed changes in diagnostic criteria. *Journal of learning disabilities*, 46(1), 5–25. <https://doi.org/10.1177/0022219412464341>
- Tavşancıl, E. (2014). *Tutumların ölçülmesi ve SPSS ile veri analizi [Measuring attitudes and data analysis with SPSS]* (5th ed.). Nobel.
- Tosto, M. G., Momi, S. K., Asherson, P., & Malki, K. (2015). A systematic review of attention deficit hyperactivity disorder (ADHD) and mathematical ability: Current findings and future implications. *BMC medicine*, 13, 1–14. <https://doi.org/10.1186/s12916-015-0414-4>
- Uçar, S., & Duman, G. (2024). *Erken Çocuklukta Oyun Temelli Etkinliklerle Diskalkuli-Matematik Öğrenme Güçlüğüne Müdahale*. Eğiten Kitap.
- Von Aster, M. G., & Shalev, R. S. (2007). Number development and developmental dyscalculia. *Developmental medicine & child neurology*, 49(11), 868–873. <https://doi.org/10.1111/j.1469-8749.2007.00868.x>
- Wheaton, B., Muthen, B., Alwin, D. F., & Summers, G. F. (1977). Assessing reliability and stability in panel models. *Sociological Methodology*, 8, 84–136. <https://doi.org/10.2307/270754>
- Yoong, S. M., Ahmad, N. A., Swaran Singh, C. K., & Wong, W. L. (2023). The design and development of a dyscalculia checklist based on a focus group interview. *British Journal of Special Education*, 50(3), 403–412.
- Zahra, J., Jamil, F., & Khalid, R. (2014). Development and preliminary validation of an indigenous scale for assessment of learning disabilities. *Pakistan Journal of Social and Clinical Psychology*, 12(2), 27.
- Zhang, S., Liu, J., Wang, J., Xia, X., Zhang, L., Liu, L., & Jiang, T. (2019). Developing and validating the Learning Disabilities Screening Scale in Chinese elementary schools. *International Journal of Educational Research*, 96, 91–99.

APPENDIX

English and Turkish versions of the scale

Items	Never	Rarely	Sometimes	Usually	Always
Number Sense					
36. Inverts or changes the order of numbers when reading and writing (for example, writes or reads 69 as 96 or 342 as 234)					
38. Writes numbers to the wrong digit in mathematical operations					
35. Difficulties in determining the place value (especially numbers with zero tens digit in three-digit numbers (for example, 305))					
40. Difficulties in establishing a one-to-one relationship between number symbols and objects when counting					
39. Has difficulty rounding numbers to tens or hundreds (especially numbers with 5 in ones and tens digits)					
43. Difficulties in rhythmic counting backwards starting from any number					
47. Difficulties in comparing numbers (e.g. less, more)					
32. Difficulty in showing the position of any number on the number line					
Calculation					
26. Has difficulty in calculating the total amount when shopping					
27. Has difficulty in calculating change when paying for shopping					
28. Has difficulty in reading an analogue clock					
Mathematical Reasoning					
6. Has difficulty in using different problem solving strategies when solving problems					
4. Difficulties in interpreting mathematical problems					
17. Difficulties in mathematical thinking					
10. Difficulties in interpreting numerical data					
19. Difficulties in concretising and visualising abstract mathematical concepts (modelling)					

Maddeler	Hiçbir zaman	Ara sıra	Bazen	Genellikle	Her zaman
Sayı Hissi					
36. Sayıları okurken ve yazarken ters çevirir veya sırasını değiştirir (örneğin, 69 sayısını 96 veya 342 sayısını 234 olarak yazar veya okur)					
38. Matematiksel işlemlerde sayıları yanlış basamağa yazar					
35. Basamak değerini belirlemede zorluk yaşar (Özellikle üç basamaklı sayılarda onlar basamağı sıfır olan sayılar (örneğin, 305))					
40. Sayma yaparken sayı sembolleri ile nesneler arasında birebir ilişki kurmakta zorlanır					
39. Sayıları onluğa veya yüzlüğe yuvarlamakta zorluk yaşar (Özellikle birler ve onlar basamağında 5 olan sayılar)					
43. Herhangi bir sayıdan başlayarak geriye doğru ritmik saymada zorlanır					
47. Sayıları karşılaştırmalarda güçlük çeker (örn. daha az, daha fazla)					
32. Sayı doğrusu üzerinde herhangi bir sayının konumunu göstermekte zorlanır					
Hesaplama					
26. Alışveriş yaparken toplam tutarı hesaplamakta zorluk yaşar					
27. Alışverişte ödeme yaparken para üstünü hesaplamada zorluk yaşar					
28. Analog saati okumakta zorluk yaşar					
Matematiksel Akıl Yürütme					
6. Problem çözerken farklı problem çözme stratejilerini kullanmakta zorluk yaşar					
4. Matematiksel problemleri yorumlamada güçlük yaşar					
17. Matematiksel düşünmede zorluk yaşar					
10. Sayısal verileri yorumlamakta güçlük yaşar					
19. Soyut matematiksel kavramların somutlaştırılması ve görselleştirilmesinde (modelleme) zorluk yaşar					



| Research Article / Araştırma Makalesi |

Level of Middle School Students' Achieving Objectives Common with Developed Countries in Mathematics

Matematik Alanında Gelişmiş Ülkelerle Ortak Olan Hedeflere Ortaokul Öğrencilerinin Ulaşma Düzeyi¹

Merve BİLGİÇ KULAK², Özge CAN ARAN³

Keywords

1. Mathematics
2. Curriculum
3. Developed Countries
4. Middle School
5. Comparative Study

Anahtar Kelimeler

1. Matematik
2. Öğretim Programı
3. Gelişmiş Ülkeler
4. Ortaokul
5. Karşılaştırmalı Araştırma

Received/Başvuru Tarihi

19.04.2025

Accepted / Kabul Tarihi

10.10.2025

Abstract

Purpose: This study aims to determine the achieving level of eighth grade students in Türkiye to mathematics objectives common with the developed countries (namely the USA, Canada, Japan and Germany) in Mathematics, and the factors affecting this level.

Design/Methodology/Approach: Descriptive research design was used in this study. The sample of this study consisted of 538 eighth graders and 15 mathematics teachers in middle schools located in central districts with high socioeconomic status in capital city of Türkiye. In this study, curriculum analysis form, mathematics achievement test (KR 20=.86) and teacher interview form were used as data collection tools.

Findings: Results indicated that some of the mathematical skills, learning areas and objectives were different in curricula, the achieving level of common objectives by students is low (63%) and Turkish teacher's opinions on mathematics course pointed about the arrangements for the countries that aim to succeed in mathematics.

Highlights: There are similarities and differences in the mathematics curricula of developed countries and Türkiye. The mathematics achievement of students attending schools in the capital city of Türkiye, which has a centralized education system, is below the expected level. Mathematics teachers state that arrangements should be made in the curriculum and implementation of it.

Öz

Çalışmanın amacı: Bu araştırmanın amacı, Türkiye'deki sekizinci sınıf öğrencilerinin, gelişmiş ülkelerle (ABD, Kanada, Japonya ve Almanya) ortak matematik hedeflerine ulaşma düzeylerini ve bu düzeyi etkileyen faktörleri belirlemektir.

Materyal ve Yöntem: Bu çalışmada betimsel araştırma deseni kullanılmıştır. Bu çalışmanın örneklemini Türkiye'nin başkentinde sosyoekonomik düzeyi yüksek merkez ilçelerdeki ortaokullarda öğrenim gören 538 sekizinci sınıf öğrencisi ve 15 matematik öğretmeni oluşturmaktadır. Araştırmada veri toplama aracı olarak program analiz formu, matematik başarı testi (KR 20=.86) ve öğretmen görüşme formu kullanılmıştır.

Bulgular: Sonuçlar, öğretim programlarında bazı matematiksel beceri, öğrenme alanı ve kazanımların farklı olduğunu, öğrencilerin ortak kazanımlara ulaşma düzeylerinin düşük olduğunu (%63) ve Türk öğretmenlerinin matematik dersine ilişkin görüşlerinin, matematikte başarılı olmayı hedefleyen ülkelere yönelik düzenlemelere işaret ettiğini göstermektedir.

Önemli Vurgular: Matematik alanında gelişmiş ülkelerin ve Türkiye'nin matematik dersi öğretim programlarında farklılaşan ve benzeyen noktalar vardır. Merkezi eğitim sistemine sahip Türkiye'nin başkentindeki okullara devam eden öğrencilerin matematik başarıları beklenen düzeyin altındadır. Matematik öğretmenleri matematik dersinin öğretim programı ve uygulanma sürecinde düzenlemeler yapılması gerektiğini belirtmektedirler.

¹ This study is based on the first author's master thesis submitted to Hacettepe University Graduate School of Educational Sciences under the supervision of the second author.

² Mevlâna Middle School, Düzce, Türkiye; <https://orcid.org/0000-0002-5772-6115>

³ **Corresponding Author**, Hacettepe University, Faculty of Education, Department of Curriculum and Instruction, Ankara, Türkiye; <https://orcid.org/0000-0003-3229-4325>

INTRODUCTION

The changes that have occurred in our age and the changing needs of the individual and society have also directly affected the roles expected from individuals. Therefore, raising individuals who are thinking critically, solving problems, determined, enterprising, empathetic, having communication skills and, contributing society and culture...etc. has become important as a result of this change (MoNE, 2018). In addition, the individual's competence in mathematics, science, and technology (MONE, 2018) is also seen among the important characteristics today. In this context, with the qualified education that countries offer in basic disciplines, it is possible to meet the qualified manpower needs of a country by ensuring development of individual who could keep up with the age. At this point, Aksu (2008) emphasized that the quality of mathematics education should be constantly increased in order to raise qualified individuals.

Mathematics has been used by many societies from past to present and has always been considered important in the process of human development and the progress of societies (Kahramaner and Kahramaner, 2003). Many countries included in the group of developed countries attach importance to mathematics as a basic discipline in order to ensure their progress (TMA, 2016). In addition, it would be misleading to talk about science and technology, a qualified product and service, and socio-economic development without mathematics (Ersoy, 2003). Since mathematics occupies such an important place in our lives, sheds light on different scientific fields, and contributes to the development of societies, the teaching of mathematics and the training of individuals competent in the field of mathematics in order for countries to reach the level of contemporary civilization should be given importance. In this context, it is important to examine how mathematics is taught in developed countries, and the results of this analysis and also international mathematics exam or ranking will shed light on researchers in terms of organizing mathematics curriculums of countries that lag behind in international rankings in the field of mathematics education. Because curriculums guide teachers in the learning-teaching process. In curriculums, the basic knowledge and skills that students are expected to gain at the end of the learning-teaching processes are prepared by separating them into class levels. Evaluating the prepared and implemented curriculums is important in terms of determining whether students reach the desired goals as a result of the teaching, whether students gain the nature of knowledge, the deficiencies of the programs, and the problems that arise during the implementation of the program in the learning-teaching processes. When evaluating curriculums, these evaluations must be made according to certain criteria. At this point, examining the curriculums of developed countries in education of a discipline to determine criteria and analysing the strengths and weaknesses of a curriculum belongs to developing countries in that discipline, will help countries that lag behind in the education of that discipline to improve themselves .

Türkiye participates in some international exams (such as PISA and TIMMS), International Mathematical Olympiads (IMO) and associations such as International Mathematics Union (IMU), which rank countries according to their success in mathematics, in order to determine where it stands compared to other countries and where it is lacking. Considering the information in the reports of various organizations and associations that measure certain skills such as whether students have acquired the nature of knowledge, whether they produce rational solutions to problems encountered in daily life, and organize exams to rank countries' success in mathematics according to these criteria, it was deemed necessary to conduct such a study in order to provide an effective education that will help students in Türkiye grow up in accordance with the needs of the age. In this context, this study aimed to examine the mathematics curriculum of Türkiye and developed countries in the mathematics education. The purpose of the analysing the curricula was to reveal to what extent the common goals that developed countries expected to gain in mathematics until high school, are achieved through the mathematics curriculum in Türkiye and to obtain teachers' opinions on the factors that affect the level of students in Türkiye achieving these common objectives.

This research was grounded in curriculum evaluation of developed countries in mathematics and Türkiye. Curriculum evaluation inform the current level of success in line with the achievable goals, the effectiveness of the program in achieving the goals, and the weak and strong points of the curriculum (Doğan, 1997; Güven and İleri, 2007; Ornstein and Hunkins, 2009; Tyler, 2014; Varış, 1994). In most of the comparative studies about program evaluation, the curricula of the countries were compared with document analysis (e.g. Abid, 2017; Bacakoğlu and Tertemiz, 2021; Çetinbağ, 2019; Erbilge, 2019; Güzel, 2010; Kaytan, 2007; Özkan, 2006; Sugandi, 2015; Vezina, 2023; Wilms, 2011). According to Demirel (2019), seeking solutions to the educational problems at school will provide more data. At this point, the data obtained from schools is crucial in evaluating the curriculum used for students to achieve various characteristics. Therefore, necessary improvements in curriculum, based on the scientific data obtained from schools will (Varış, 1994), contribute to students' development more. At this point, this research aims to answer following questions:

- (1) What are the common and different characteristics (including mathematical skills, learning areas and objectives) emphasized in secondary school mathematics curriculum of developed countries in the Mathematics and Türkiye?
- (2) To what extent do students in Türkiye achieve the common objectives included in the mathematics curricula? and,
- (3) What do teachers in Türkiye think about the factors affecting students' achieving common objectives?

In the literature, there are studies measuring the mathematics-related skills of middle school students in Türkiye. The results of these studies indicate that students' math skills are not at the expected level. For example, according to the research results of Uysal and Yenilmez (2011), it was concluded that the vast majority of students are at low levels in terms of mathematical literacy. Çakırtekin (2025) also found that students had an average level of performance in mathematics literacy problems and that their strategic flexibility was low. Sönmez (2025) also concluded that students' mathematics literacy proficiency levels were at or below

the minimum level. The research results of Alabayrak, Tarım, and Baypınar (2023) also showed that gifted students' mathematics literacy achievements were average. Unlike previous studies in the literature aimed at determining middle school students' mathematics literacy skills, this study comprehensively sought to determine the extent to which eighth-grade middle school students in Türkiye achieve internationally expected targets in mathematics. Furthermore, unlike many studies in the literature that only perform comparative program analysis through document analysis, this study also collected data from students and the teachers implementing the program. The results of this study are expected to inform the arrangement by both the countries that are behind in terms of mathematics success, such as Türkiye, and developed but aim to continue their success in mathematics.

METHOD/MATERIALS

Descriptive method was used in this study to determine the level of eighth grade Turkish students' achieving mathematics objectives common with the developed countries in Mathematics, and the factors affecting this level. The descriptive method describes the details of the characteristics of individuals, groups or physical environments in detail (Büyüköztürk et al., 2019). In this context, quantitative and qualitative data collection methods were used together in the study.

Study Group

The data sources of this study namely documents, students and teachers, were selected by criterion and maximum diversity sampling technique. While choosing mathematics curricula as document to analyse in this research, ranking among countries participated in International Mathematical Union 2018 (IMU, 2018), PISA 2018 (MoNE, 2019), TIMSS 2015 (MoNE, 2016a), International Mathematical Olympiad 2021 (IMO, 2023), is used as criteria. At this context, the curricula of successful countries in mathematics, namely ABD (California), Canada (Ontario), Japan, Germany (Berlin), and Türkiye as a country left behind in ranking were chosen. The curricula of these countries (Blum et al., 2010; California, 2015; Common Core, 2020; Elementary, 2008; Junior, 2010; Mathematik, 2017; The Ontario, 2005) and Türkiye (MoNE, 2018) were accessed through the documents published on the official websites of the Ministries of Education. In addition, this research is conducted with 538 eighth grade students and 15 mathematics teachers in secondary schools located in high socioeconomic areas of central districts of the capital city of Türkiye. School's being located in the area which is advantageous in terms of geographical situation, economic and social development level, transportation conditions and meeting service requirements (MoNE, 2016b) was used as criteria and maximum variety was provided by choosing school from each central districts of the capital city of Türkiye (8 districts total in the capital). International exam results (e.g. PISA, TIMSS) reveal that there is a directly proportional relationship between the quality of schools and student success (MoNE, 2023; OECD, 2004; OECD, 2005; OECD, 2008 and OECD, 2019; OECD, 2023). The fact that each school in Türkiye is different in terms of success is associated with the different socioeconomic levels of students (Özer and Suna, 2021). In this context, it was aimed to reach students and teachers of the schools with qualified educational conditions in order to provide a rich data source for the research.

Data Collection

Data were collected by document analysis form, mathematics achievement test and interview form in this study after obtaining the approval from Ethics Committee and Ministry of National Education for the implementation of the research. In the study, firstly, the middle school mathematics course curricula of developed countries and Türkiye were examined according to categories (included mathematical skills, learning areas, subjects, objectives) in the curriculum review form developed by the researcher, and the common and different characteristics that these countries expected to provide students before high school in mathematics course, were determined.

For the content validity of the curriculum review form, the opinions of two program development experts and one mathematics education expert were consulted. The form was designed according to the expert opinions. The mathematics course curricula of developed countries to be examined through the curriculum review form.

The achievement levels of eighth grade students in Türkiye to the common goals determined within the scope of the research were tried to be determined with a mathematics achievement test consisting of real-life problems developed by the researcher (see Appendix 1 for question sample). The opinions of two measurement and evaluation experts (one of them is both a measurement and evaluation and mathematics education expert), three program development experts (two of them are both program development and mathematics education experts) and two mathematics education experts were consulted in terms of the suitability of each item in the mathematics achievement test to measure the determined target behaviours, whether the items related to these target behaviours were written in accordance with the Bloom Taxonomy levels, whether the items are in accordance with the multiple-choice item writing rules, scientific nature and suitability of the items to the student level. As a result of the expert opinions, necessary corrections were made in the test created and some items were removed from the test. In order to test the comprehensibility of the test items in terms of language and expression, all the items were read by three eighth grade students. Concept and expression corrections were made in the items that were not understood by the students. The pilot application form was applied to 80 eighth grade students. The data obtained from the pilot application were analysed using the TAP (Test Analysis Program), and item discrimination, item difficulty indices, and the KR-20 reliability coefficient of the test were calculated. When selecting test items, items with an item discrimination of over 0.20 and items with an item difficulty index of

medium difficulty (over 0.30) were preferred (Özçelik, 1989; Büyüköztürk et al. 2019). The choice analyses of the items were also checked, distractors were strengthened, and ineffective choices were developed. Based on these analyses, problematic items were removed from the test, and items that needed correction were improved. The KR-20 reliability coefficient of the final form (74 items) was calculated as 0.86. The Mathematics Achievement Test was administered to 538 volunteer students, including the pilot application and the main application, by having the student and parent consent forms signed. The data collected as a result of the main application were re-analysed with the TAP (Test Analysis Program) and the KR-20 reliability coefficient of the test was calculated as 0.94.

In order to obtain teachers' opinions on the factors affecting the level of eighth grade students' achieving these common objectives within the scope of the mathematics course curriculum used in our country, a teacher interview form was created by the researcher. For the content validity of the teacher interview form, a total of six experts were consulted, including one measurement and evaluation expert, three program development experts (two of whom are both program development experts and mathematics education experts), and two mathematics education experts. The opinions of two Turkish teachers were also obtained regarding the conformity of the interview questions in the form to the language-expression rules. In line with the expert opinions, some interview questions were corrected, and some questions were grouped under a single question as probes and expression errors were eliminated. The draft form was applied to three mathematics teachers who were not included in the research sample, and the clarity of the interview questions was tested, and the necessary corrections were made according to the opinions received from these teachers, and the final form consisting of six open-ended questions was created. During the application process of the interview form, 15 teachers in the sample signed a consent form. The answers given by the volunteer teachers to the interview questions were audio-recorded with their permission in order to avoid data loss.

The qualitative data were collected by document analysis and interview. In this process, for accurate representation of reality through research results, the mathematics curricula used in this research were translated into Turkish with three different translation programs and before the translations were made, the programs were tested with English teachers and the translation results were checked. Also analyzing curricula with the same document analysis form provided systematic data collection and contributed to the reliability of data. In addition, to ensure the credibility of the data obtained from the research, attention was paid to asking the interview questions in a clear way and "reading aloud" was performed in three volunteer teachers who would not participate in the final study. To ensure the transferability of the obtained data, the teachers were asked to explain their thoughts in detail and give examples.

Data Analysis

The quantitative data of the research were analysed with descriptive statistics. In this process, the percentage of correct answers given by eighth grade students to 74 items in the applied mathematics achievement test was examined. Percentages were calculated and examined separately according to each learning area, subject content and cognitive area steps (knowledge, comprehension, application and analysis).

Besides the qualitative data obtained from documents and interviews were analysed with a deductive approach involves coming to the data with some preconceived themes. In this context, the common and different characteristics in the mathematics curricula were examined through document analysis and presented in tabular form. In the analysis of the qualitative data obtained from the interviews, first of all, the audio recordings obtained during the interview were listened to twice and turned into a written document one by one by hand. Then, coding was done by using the dimensions in the interview form (characteristics of the goals, organization of the content, teaching- learning and evaluation process), in other words, the previously determined themes. Coding helps to identify and label various aspects of the data, and is a structure that separates the data converted into written documents into meaningful pieces and maintains the integrity of meaning among these pieces (Corbin and Strauss, 2008; Creswell, 2013; Merriam, 2009; Miles and Huberman, 2016), while themes are broad units of information that form codes for similar views (Corbin and Strauss, 2008; Merriam, 2009). The qualitative data obtained from the interviews were presented by percentage, and quotations were used for the validity of the qualitative data. While presenting the quotations, teachers were coded as T1, T2, and T3. Inter-coder reliability was also calculated for the data obtained from the interview after the data was analyzed by researchers and an expert read and coded data again. The codes determined by the expert were compared with the codes of the researcher while calculating inter-coder reliability. The agreement rate between the researcher and the expert in the coding of the data obtained from the interview form was determined as 81.25%. This rate being 70% and above means that the research is considered reliable in terms of data analysis (Miles and Huberman, 2016). Then, experts discuss the inconsistent points in the matches and decipher them. They reached a common conclusion regarding the data-theme match.

FINDINGS AND DISCUSSION

In order to answer research questions, first of all, similarity and differences in the secondary school mathematics curriculum of developed countries in Mathematics and Türkiye were defined. The common and different skills of Türkiye and developed countries in the field of mathematics are summarized in Table 1.

Table 1. The Distribution of Mathematical Skills by Countries

	United States	Germany	Japon	Canada	Türkiye
Problem Solving	x	x	x	x	x
Research	x	x	x	x	x
Reasoning	x	x	x	x	x
Reflection		x		x	
Mathematical modelling	x	x		x	
Making connections (association))	x	x	x	x	x
Mathematical communication (Using symbolic, formal and technical elements of mathematics)		x		x	x
Abstract thinking	x	x	x	x	x
Relating mathematics to daily life	x		x	x	x
Taking pleasure in mathematics		x	x		x
Ability to use mathematical tools and choose computational strategies			x	x	
Persistence in mathematics					x

Table 1 shows that the common skills related with mathematics emphasized in the mathematics curricula were problem solving, research, reasoning, connection and abstract thinking. On the other hand, the skills not common in the mathematics curricula of all countries, are reflecting (only in Germany and Canada), mathematical modeling (only in the USA, Germany and Canada), mathematical communication (only in Germany, Canada and Türkiye), associating mathematics with daily life (in all countries except Germany), taking pleasure in mathematics (only in Germany, Japan and Türkiye), using mathematical tools and choosing calculation strategies (only in Japan and Canada), and showing persistence in mathematics (only in Türkiye).

The document analysis of mathematics curricula also indicated that there are four common learning areas in the curricula. These are Numbers and Operations, Geometry and Measurement, Algebra, Data Analysis and Probability. The curriculum analysis also showed that the "Probability" learning area is included in the "Use of Data" learning area in the Japanese mathematics curriculum, and the "Mathematical Relationships" learning area is also included in the Japanese mathematics curriculum. In addition, Canada's curriculum, unlike other programs, includes the "Pattern" learning area. Another result of curricula analysis is about the contents and it showed that the contents are mostly similar, but the grade levels in which the subjects were taught differ. Therefore, the subjects different, were Interest and Binomial in Germany, Function in Germany and Japan, Sets, Square Root Expressions and Views of Objects from Different Directions in Türkiye, Transformation Geometry in the USA, Türkiye and Canada, Pythagorean Theorem in Canada, Germany and Türkiye and Folded Shapes in Japan.

In this research, the objectives related with the subjects in four common learning areas were also examined and this analysis showed that a total of 74 target behavioral objectives related with Integers, Divisors and Multiples, Fractions, Exponents, Decimal Numbers, Percentages, Ratio-Proportion, Geometric Shapes, Measurement, Circle and Circle, Algebraic Expressions, Data Analysis, are common. Also, the behavioral objectives at the analysis level of the curricula according to Bloom Taxonomy (Anderson et al.2018; Bloom,1956), were analyzed and the results indicated that "drawing the expansion of 3-dimensional shapes" in Türkiye and the USA, and "finding the general rule of a pattern" in Türkiye, the USA and Canada were common whereas "statistically analyzing data presented in tables or graphs" is common in all countries.

After the document analysis, the level of Turkish students' achieving 74 behavioral objectives common with developed countries in Mathematics was determined as 63% by the help of achievement test and this level is not at the expected level. Because students are expected to achieve 80-95% level of learning of the target behaviors in progressive courses such as mathematics (Özçelik, 1981; Yıldırım, 2006) before moving on to high school, in order to equipped with the mathematical skills necessary for being ready for the global world when it is considered that cognitive entry behaviors of a course can be explained

up to half of the variability in success($r=0.70$) in subsequent learning units (Bloom,2016). The distribution of levels of achieving objectives by learning areas are also determined in this research. Figure 1 shows this distribution.

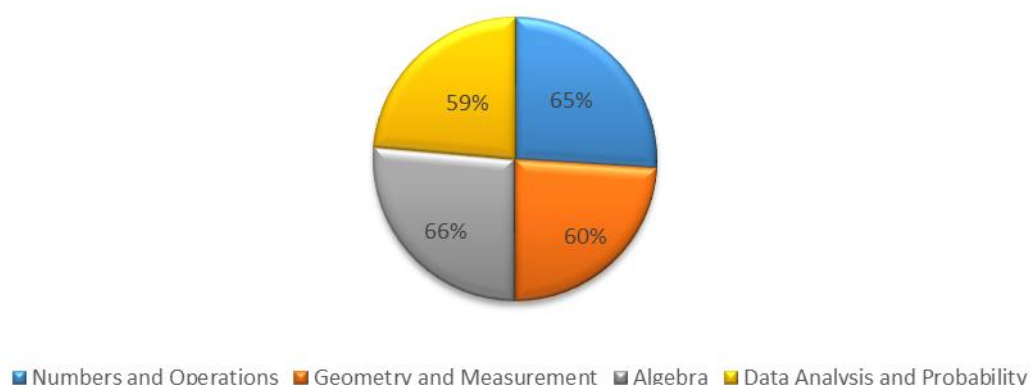


Figure 1 The levels of achieving objectives in the learning areas

When Figure 1 is examined, it is seen that the level of achievement of target behaviors in the Numbers and Operations learning area is 65%, the level of achievement of target behaviors in the Geometry and Measurement learning area is 60%, the level of achievement of target behaviors in the Algebra learning area is 66% and the level of achievement of target behaviors in the Data and Probability learning area is 59%. It was determined that the highest target behaviors were achieved in the Algebra learning area and the lowest target behaviors were achieved in the Data and Probability learning area. When the level of achievement of target behaviors in each learning area was examined; it was determined that the target behaviors were not achieved at the desired level (75%). This situation is consistent with studies in the literature that reveal students' low levels of mathematical literacy (Sönmez, 2025; Uysal and Yenilmez, 2011). Çakırtekin's (2025) study is consistent with these research findings in some respects, revealing that even students with an average level of mathematical literacy perform poorly in sub-skills such as strategic flexibility. Albayrak, Tarım, and Baypınar (2023)' study showing that even gifted students demonstrate moderate mathematical literacy skills also corroborates the findings of this research in some respects. The distribution of common objectives in the mathematics curriculum of developed countries and Türkiye according to Bloom taxonomy levels is also determined in this research. Table 2 shows this data.

Tablo 2. The Levels of Achievement of All Common Goals That Developed Countries Want to Achieve in Mathematics Classes by 8th Grade Students in Türkiye

Levels of Items in the Achievement Test According to Bloom Taxonomy	The number of common objectives in each level of Bloom Taxonomy	Level of Achieving Objectives
Knowledge Level	21	% 64
Comprehension Level	22	% 66
Application Level	28	% 60
Analysis Level	3	% 71

When Table 2 is examined, according to the findings obtained from the applied mathematics achievement test, it is seen that the eighth-grade students could not exceed the 75% level determined as the criterion for the target behaviors at the knowledge, comprehension, application and analysis levels included in the test in general. It was determined that the levels of reaching the target behaviors at the knowledge, comprehension and application levels were close to each other. It was determined that the level of reaching the three target behaviors at the analysis level was higher than the other target levels.

In addition, the reason why students did not achieve the objectives in expected level, was also revealed in this research. The qualitative data obtained from teacher interviews gave detailed information related with students' achieving level. At this point, about general structure about curriculum, teacher stated that the number of objectives in the curriculum is high (80%), curriculum does not provide flexibility to the teacher in the teaching- learning process (67%), the spiral curriculum approach is not applied correctly (33%), the curriculum did not allow the student to connect with daily life (13%) and the curriculum was not learner-based (13%). Also, teacher's opinions about objectives noted the reasons. These are the behavioral objectives' including more than one behavior (60%), and only focusing to achieve the limited mathematical skills such as problem solving, reasoning and association skills (40%). In this context, some of the teachers' opinions explaining the reasons why students do not reach the desired level are as follows:

"I find the curriculum inadequate in general. I cannot relate the student to daily life. This is easier in some grades. The outcomes are written in such a way that there are too many. It should definitely be reduced. Students are not happy with being bombarded with so much information, they have difficulty in. Time is limited but the outcomes are many. When we try to do a lot of work in a short time, we have to move on to another before the student has mastered one subject." (T2)

"The Ministry of National Education says that the ideology of the curriculum is learner-based, but it is not a program that can include students in the process. Unfortunately, it is inconsistent in this regard. There are too many objectives. Students complain about being taught the same topics over and over again. For example, exponential numbers were taught for consecutive years. You can explain it in detail once or twice and move on. It does not seem right to spread some topics to all levels. More time can be given to other topics." (T3)

Another notable finding of the study is that one of the teachers stated that the mathematics curriculum only prepared students for the analysis level:

"It supports skills that can achieve the maximum level of analysis. The program is not appropriate for the student's development level. There is still a lot of content that is unnecessary to use in middle school and is difficult for the student. I think it should be simplified. At least the objectives can be simplified. We expect the student to do too much. When they can't do it, we consider them unsuccessful." (T1)

According to teachers' opinions regarding the teaching-learning process, it is not possible to connect the subjects with daily life due to standard exam preparation in the upper levels of secondary school (7th and 8th grade) (33%) and a limited number of activities are carried out due to the limited teaching-learning time (20%).

"In the lower grades, connections can be made to daily life. For example, I can divide an apple in the subject of fractions, but this is not possible in the upper grades. Since we are exam-focused, we worry about catching up on topics and solving questions, or even solving new generation questions. We cannot do activities. In the upper grades, it works a bit more like private teaching institution. No matter how many questions I solve, it is not enough. Therefore, it is not possible to make connections with other disciplines. There is no time for this. As I said in the first question. There are many objectives." (T15)

Teachers' opinions about textbooks also indicated that the question patterns that will measure the characteristics expected to be acquired by students are insufficient (67%), the books do not contribute to higher order thinking skills (47%), and there are errors in the contents (20%).

"The books are appropriate for the student's level, but they do not contain questions or content that will prepare them for the high school entrance (LGS) exams. They mostly contain questions that are sufficient to allow us to determine the student's learning level throughout the process. I also do not think they contribute to higher-level skills. We do not offer students a program that will develop this skill so that we can see this in the books." (T8)

Teachers' opinions regarding measurement and evaluation methods showed that the methods direct the student to memorization, knowing enough to pass the exam, and anxiety about passing the exam (40%), and are not aimed at measuring the student's higher order thinking skills (33%):

"I try to apply summative tests, comprehension tests, subject achievement tests to the students. However, the exams are mostly in the form of students writing what they know and seeing what they have learned and not learned, so they are monotonous traditional exams. For this reason, I cannot do an application that will measure high-level skills. When we do such applications, if the student gets a low score, the parent immediately asks why my child got a low score. When we do a traditional exam, the student's exam results are better. The student may also be memorizing, cheating is inevitable anyway. I do open-ended exams, at least I can prevent cheating compared to the test exam." (T8)

According to teachers' opinions, students are not able to learn the subjects meaningfully due to reasons such as not being able to allocate enough time to important subjects or activities because the mathematics curriculum contains too many objectives, the objectives being overlapping and the spiral curriculum approach not being applied correctly, subject's not being sufficiently transferred to daily life, and the curriculum not being student-centered. In addition, content does not develop high-level thinking skills and measurement and evaluation activities direct the student to memorize information instead of meaningful learning. Although Türkiye's mathematics curriculum addresses many mathematical skills, similar to developed countries, there are limited objectives pointed some of these skills in the mathematic curriculum. This can cause the inconsistency between the official curriculum and operational curriculum (Posner, 1995) and all these reasons stated by teachers noted why Türkiye is fall behind other countries in mathematics. In progressive courses such as mathematics, it is not possible for students to acquire objectives without fully achieving others. At the same time, if the activities carried out in the teaching-learning process, content and measurement and evaluation activities, in other words, quality of instruction is not improved, it is not possible for the students to reach the expected learning level (Bloom, 2016).

CONCLUSION AND RECOMMENDATIONS

This research aimed to analyze the level of eighth grade Turkish students' achieving mathematics objectives common with the developed countries in Mathematics, and the factors affecting this level and the results showed that students in Türkiye, have not

been able to achieve objectives common with developed countries in mathematics at the expected level. Teacher opinions have revealed the reasons for this situation. After this research were done, mathematics curriculum of Türkiye was just changed. So, the results of this research are limited with the data about Curriculum developed in the year 2019. Considering that the new mathematics curriculum developed in 2024 is not applied in eight grade level yet and curriculum development is a process, the data of this research can guide the improvements expected to be done in mathematics course. Also, collecting data from the capital city can be thought as a limitation but centralized education system is applied in Türkiye and the data of the research reflects the schools' being located in the area which is advantageous and maximum variation ensure external validity (transferability), was considered while schools were chosen. In addition, the findings of this research conducted on the Turkish sample, are very important for the countries which aim to develop their mathematic education. The qualitative and quantitative data of this research draw attention to the skills, objectives and content that should be given importance in order to increase the effectiveness of mathematics curriculum, and draw attention to the priority points that program developers and teachers as program implementers, should focus on to ensure that students reach the goals at the desired level in mathematics.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

This research did not receive any funding

Statements of publication ethics

We hereby declare that the study has not unethical issues and that research and publication ethics have been observed carefully.

Researchers' contribution rate

The study was conducted with equal collaboration of the researchers.

Ethics Committee Approval Information

The study is derived from the first author's master thesis. Ethics committee approval for the study was obtained at the Hacettepe University Ethics Commission meeting dated 14.09.2021 [Decision dated and numbered: 06.09.2021 and E-51944218-300-00001743431].

REFERENCES

- Abid, A. A. O. (2017). *İlköğretim matematik öğretim programlarının karşılaştırılması: Türkiye ve Libya (A Comparison Between The Middle School Mathematics Curriculum "A Case Study Of Turkey And Libya)* [Master's thesis, Kastamonu University].
- Aksu, H.H. (2008). Öğretmenlerin yeni ilköğretim matematik programına ilişkin görüşleri. *Abant İzzet Baysal Üniversitesi Eğitim Fakültesi Dergisi*, 8(1), 1-11.
- Albayrak, H. B., Tarım, K., & Baypınar, K. (2023). Özel Yetenekli Öğrencilerin Matematik Okuryazarlığı Öz-Yeterlik Algıları ile Matematik Okuryazarlığı Başarılarının İncelenmesi. *Trakya Eğitim Dergisi*, 13(1), 115-127.
- Anderson, Lorin W., Krathwohl, David R., Airasian, Peter W., Cruikshank, Kathleen A., Mayer, Richard E., Pintrich, Paul R., Raths, J. & Wittrock, Merlin C. (2018). *Taxonomy for learning, teaching, and assessing, a: a revision of bloom's taxonomy of educational objectives* (D.A. Özçelik, Trans.). Pegem Akademi.
- Bacakoğlu, T.Y. & Terteimiz, N. (2021). Türkiye ve singapur ilköğretim matematik dersi öğretim programlarının geometri öğrenme alanı bağlamında karşılaştırılması (A comparison of mathematics curriculum of Türkiye and Singapore in the context of geometry learning area). *Nevşehir Hacı Bektaş Veli Üniversitesi Dergisi*, 11(3), 1089-1107.
- Bloom, B. S. (2016). *Human Characteristics and School Learning* (D. A. Özçelik, Trans.). Pegem Yayıncılık.
- Bloom, B. S. (1956). *Taxonomy of educational objectives, handbook 1: Cognitive domain*. Longman.
- Blum, W., Drüke-Noe, C., Hartung, R. ve Köller, O. (2010). *Bildungsstandards Mathematik: konkret*. <https://edoc.hu-berlin.de/bitstream/handle/18452/3776/4.pdf?sequence=1>
- Büyüköztürk, Ş., Çakmak, E., Akgün, Ö. Karadeniz, Ş. & Demirel, F. (2019). *Eğitimde bilimsel araştırma yöntemleri*. Pegem Akademi
- Çakırtekin, M. (2025). *Ortaokul öğrencilerinin matematik okuryazarlığı problemlerini çözme performansları ve stratejik esnekliklerinin incelenmesi (Mathematical Literacy of Middle School Students Problem Solving Performance And. Examining Their Strategic Flexibility)*. [Master's thesis, Uludağ University].
- California Department of Education. (2015). *Overview of the standards chapters of the mathematics framework for California public schools: Kindergarten through grade twelve*. <http://www.cde.ca.gov/ci/ma/cf/documents/mathfwoverview.pdf>
- Common Core States Standards for Mathematic. (2020). <https://learning.ccsso.org/wp-content/uploads/2022/11/ADA-Compliant-Math-Standards.pdf>
- Corbin, J. & Strauss, A. (2008). *Basics of qualitative research techniques and procedures for developing grounded theory (third edition)*. Sage.
- Creswell, J. W. (2013). *Qualitative inquiry and research design: Choosing among Five approaches*. Sage.
- Çetinbağ, A. (2019). *Türkiye ve Kanada ilköğretim matematik öğretim programlarının program öğeleri bağlamında karşılaştırılması (Comparing turkey and Canada in the context of the program elements of primary school mathematics curriculum)* [Master's thesis, Marmara University].

- Demirel, Ö. (2019). *Eğitimde program geliştirme: Kuramdan uygulamaya (Curriculum Development in Education: From Theory to Practice)*. Pegem Akademi
- Doğan, H. (1997). *Eğitimde program ve öğretim tasarımı (Curriculum and Instructional Design in Education)*. Önder Matbaacılık
- Elementary School Teaching Guide for the Japanese Course of Study: Mathematics (Grade 1-6) (2008). https://www.lsalliance.org/wp-content/uploads/2016/10/Japanese_COS_TeachingGuide_en.pdf
- Erbilge, E.A. (2019). Türkiye, Kanada ve Hong Kong'un ortaokul matematik öğretim programlarının karşılaştırılması (Comparison of secondary school mathematics curriculum of Turkey, Canada And Hong Kong) [Master's thesis, Marmara University].
- Ersoy, Y. (2003). *Teknoloji Destekli matematik eğitimi: Gelişmeler, politikalar ve stratejiler. İlköğretim-Online*, 2(1),18-27. <http://www.ilkogretimonline.org.tr/index.php/io/article/view/2055/1891>
- Güven, B., & İleri, S. (2006). Program değerlendirme kavramı ve ilköğretimde program değerlendirme çalışmalarına kuramsal bir bakış. *The Journal of Turkish Social Research*, 10(1-2), 141-163.
- Güzel, İ. (2010). *Türkiye, Almanya, Kanada ortaöğretim matematik öğretim programlarının karşılaştırmalı değerlendirilmesi (Turkey, Germany, Canada Highschool Mathematics Curriculum Comparing)* [Master's thesis, Zonguldak Karaelmas University].
- IMO [International Mathematical Olympiad] (2023). *List of countries*. <https://www.imo-official.org/countries.aspx>
- IMU [International Mathematical Union] (2018). *Statutes*. <https://www.mathunion.org>
- Junior High School Teaching Guide for the Japanese Course of Study: Mathematics (Grade 7-9) (2010). http://www.cried.tsukuba.ac.jp/math/apec/ICME12/Lesson_Study_set/Junior_high_school-teaching-guide-Mathmatics-EN.pdf
- Kahramaner, Y. ve Kahramaner, R. (2002). Üniversite Eğitiminde Matematik Düşüncesinin Önemi. *İstanbul Ticaret Üniversitesi Dergisi*. 2, 15-25. <http://acikerisim.ticaret.edu.tr/xmlui/bitstream/handle/11467/185/M00020.pdf?sequence=1&isAllowed=y>
- Kaytan, E. (2007). Türkiye, Singapur ve İngiltere matematik öğretim programlarının karşılaştırılması (Comparison of Primary Mathematics Curriculum of Turkey, Singapore and England) [Master's thesis, Hacettepe University].
- Mathematik Klasse 1-13 Übersichts, (2017). <https://www.frustfrei.lernen.de/mathematik/mathematik-klasse-uebersicht.html>.
- Merriam, S. B. (2009). *Qualitative research: A guide to design and implementation: Revised and expanded from qualitative research and case study applications in education*. Jossey-Bass.
- Miles, M. B., ve Huberman, A. M. (2016). *Qualitative Data Analysis*. (S. Akbaba Altun & A. Ersoy, Trans. Ed.). Pegem Akademi.
- MoNE (Ministry of National Education) of Türkiye (2016a). *TIMSS 2015 Ulusal Matematik ve Fen Ön Raporu (TIMSS 2015 National Mathematics and Science Preliminary Report)*. http://timss.meb.gov.tr/wpcontent/uploads/TIMSS_2015_Ulusal_Rapor.pdf
- MoNE (Ministry of National Education) of Türkiye (2016b). *Öğretmen Atama ve Yer Değiştirme Yönetmeliği Birinci Bölüm: Amaç, Kapsam, Dayanak ve Tanımlar (Teacher Appointment and Relocation Regulation Part One: Purpose, Scope, Basis and Definitions)*. https://oygm.meb.gov.tr/meb_iys_dosyalar/2017_11/08144242_mevzuat.pdf
- MoNE (Ministry of National Education) of Türkiye (2018). *İlköğretim Matematik Dersi Öğretim Programı (Primary School Mathematics Course Curriculum)*. <http://mufredat.meb.gov.tr/ProgramDetay.aspx?PID=329>
- MoNE (Ministry of National Education) of Türkiye (2019). *PISA 2018 Türkiye Ön Raporu. Eğitim Analiz ve Değerlendirme Raporları Serisi No:10 (PISA 2018 Türkiye Preliminary Report. Education Analysis and Evaluation Reports Series No: 10)*. http://pisa.meb.gov.tr/wpcontent/uploads/2020/01/PISA_2018_Turkiye_On_Raporu.pdf
- MoNE (Ministry of National Education) of Türkiye (2023). *PISA 2022 Türkiye Raporu. Sosyoekonomik Göstergeler ve Öğrenci Performansı (PISA 2022 Türkiye Report. Socioeconomic Indicators and Student Performance)*. https://pisa.meb.gov.tr/meb_iys_dosyalar/2023_12/05125555_pisa2022_rapor_051223.pdf
- Ornstein, A. C. ve Hunkins, F. P. (2009). *Curriculum: foundations, principles, and issues (Fifth Edition)*. Allyn & Bacon.
- OECD Better Life Index (2023). *Education Background*. <https://www.oecdbetterlifeindex.org/topics/education/>
- OECD. (2004). *What makes school systems perform? Seeing school systems through the prism of PISA*. OECD Publishing.
- OECD. (2005). *School factors related to quality and equity: Results from PISA 2000*. OECD Publishing.
- OECD. (2008). *Ten steps to equity in education*. OECD Publishing.
- OECD. (2019). *PISA 2018 results: Where all students can succeed (Volume II)*. OECD Publishing.
- Özçelik, D.A. (1981). *Okullarda ölçme değerlendirme (Measurement and evaluation in schools)*. ÜSYM-Eğitim Yayınları.
- Özer, M. & Suna, H.E. (2021). Türkiye'de Sosyoekonomik Düzey ve Okullar Arası Başarı Farklarının Akademik Başarı ile ilişkisi (The Relationship Between Socioeconomic Level and Achievement Differences Between Schools and Academic Achievement in Turkey). *Journal of Measurement and Evaluation in Education and Psychology*, 12(1); 54-70. <https://dergipark.org.tr/en/download/article-file/1507632>
- Özkan, E. A. (2006). *Türkiye, Belçika (Flaman) ve Singapur matematik öğretim programları üzerine karşılaştırmalı bir çalışma (A comparative study on mathematics education programs of Türkiye, Belgium (Flemish) and Singapore)* [Master's thesis, Hacettepe University].
- Posner, G.J. (1995). *Analyzing the curriculum*. US: McGraw-Hill Humanities Social
- Sönmez, D. (2025). Ortaokul 7. ve 8. Sınıf Öğrencilerinin Matematik Okuryazarlığı Yeterlik Düzeylerinin Bazı Değişkenler Açısından İncelenmesi. *Türk Eğitim Değerlendirmeleri Dergisi*, 5(2), 61-81.
- Sugandi, B. (2015). *Comparison of Turkish and Indonesian secondary mathematics curriculum; reflection of the paradigms*, [Master's thesis, Marmara University].
- The Ontario Curriculum Grades 1-8. (2005). [https://schoolweb.tdsb.on.ca/Portals/parkdaleps/docs/math18curr%20\(14\).pdf](https://schoolweb.tdsb.on.ca/Portals/parkdaleps/docs/math18curr%20(14).pdf)ISBN 0-7794-8121-6
- TMA (Turkish Mathematics Association) (2016). *Matematik raporu: Tespitler ve öneriler*. <http://tmd.org.tr/wp-content/uploads/2016/09/TMDrapor2016orijinal.pdf>
- Tyler, R.W. (2014). *Basic Principles of Curriculum and Instruction* (M. E. Rüzgar and B. Aslan,Trans). Pegem Akademi.
- Uysal, E., & Yenilmez, K. (2011). Sekizinci sınıf öğrencilerinin matematik okuryazarlığı düzeyi. *Eskişehir Osmangazi Üniversitesi Sosyal Bilimler Dergisi*, 12(2), 1-15.
- Wilms, U. H. (2011). *Comparative Study of the National Math Curriculum with Curricula from Four Nations* [Unpublished Doctoral dissertation, Piedmont College].

- Varış, F. (1994). *Eğitimde program geliştirme "teori ve teknikler"* (Curriculum development in education "theory and techniques"). Alkim Kitapçılık Yayıncılık.
- Vezina, A.C. (2023). *A Comparative Analysis of Quebec and Ontario: How Each Province Approached Elementary Provincial Numeracy Curricula and Corresponding Faculty of Education Accredited Teacher Education Programs During the Period of 2003-2018* [Unpublished Doctoral dissertation, University of Toronto].
- Yıldırım, G. (2006). *Multicultural application of mastery learning: Our thoughts,our deeds and our hopes for education*. Boğaziçi University.

APPENDIX

Appendix 1.

A real-world mathematics question sample in achievement test

The table below shows Meriam's social media usage habits.

Social Media Platform	Daily Usage Time (min)
Facebook	35
Twitter	25
Instagram	40
YouTube	45
TikTok	30

Accordingly, which of the following can be expressed as a percentage?

- A) How many minutes of social media you use per day
- B) Time belongs to Instagram of daily social media usage
- C) The most used social media platform
- D) The least used social media platform



| Research Article / Araştırma Makalesi |

A Case Study on Teachers' Knowledge and Classroom Practices about Mathematics Learning Difficulty

Öğretmenlerin Matematik Öğrenme Güçlüğüne İlişkin Bilgileri ve Sınıf İçi Uygulamaları: Durum Çalışması

Tuğba PÜRSÜN¹, Engin ÜSTÜNDAĞ²

Keywords

1. Mathematics learning difficulty
2. Teachers' views
3. Knowledge and practices
4. Qualitative research

Anahtar Kelimeler

1. Matematik öğrenme güçlüğü
2. Öğretmen görüşleri
3. Bilgi ve uygulamalar
4. Nitel araştırma

Received/Başvuru Tarihi
27.12.2024

Accepted / Kabul Tarihi
10.10.2025

Abstract

Purpose: This study aimed to determine teachers' knowledge of mathematics learning difficulty and their classroom practices.

Design/Methodology/Approach: Data were collected through a semi-structured interview form from twenty teachers selected by criterion sampling method. The content analysis technique was employed for data analysis. As a result of the analysis, two themes were identified: a) teachers' knowledge and b) teachers' classroom practices.

Findings: As a result of the research, teachers mostly defined mathematics learning difficulties as deficiencies in mathematical skills, numbers, problem solving, four operations, calculation, and difficulty in learning mathematical concepts. In classroom practices, teachers generally used lecture, learning by doing, drama, demonstration, games, peer teaching, one-to-one teaching and question-answer methods.

Highlights: The results of the study revealed teachers' knowledge and practice limitations regarding mathematics learning difficulties.

Öz

Çalışmanın amacı: Bu araştırmanın amacı öğretmenlerin matematik öğrenme güçlüğüne dair bilgilerinin ve sınıf içi uygulamalarının belirlenmesidir.

Materyal ve Yöntem: Ölçüt örnekleme yöntemi ile belirlenen yirmi öğretmenden yarı yapılandırılmış görüşme formu aracılığıyla veriler toplanmıştır. Toplanan verilerin analizinde içerik analizi tekniği kullanılmıştır. Analiz sonucunda iki tema ortaya çıkmıştır: a) öğretmenlerin bilgileri, b) öğretmenlerin sınıf içi uygulamaları.

Bulgular: Araştırma sonucunda öğretmenler çoğunlukla matematik öğrenme güçlüğüne matematik becerilerindeki eksiklikler, sayılar, problem çözme, dört işlem, hesaplama, matematiksel kavramları öğrenme güçlüğü olarak tanımlamışlardır. Sınıf içi uygulamalarda ise öğretmenlerin genellikle düz anlatım, yaparak yaşayarak öğrenme, drama, gösterip yaptırma, oyun, akran öğretimi, birebir öğretim ve soru-cevap yöntemini kullandıkları görülmüştür.

Önemli Vurgular: Araştırmadan elde edilen sonuçlar öğretmenlerin matematik öğrenme güçlüğüne dair bilgi ve uygulama sınırlılığını ortaya koymuştur.

¹ Corresponded Author, Tokat Gaziosmanpaşa University, Faculty of Education, Special Education Department, Tokat, Türkiye; <https://orcid.org/0000-0002-5436-1464>

² Vali Recep Yazıcıoğlu Primary School, Tokat, Türkiye; <https://orcid.org/0009-0007-6176-2606>

INTRODUCTION

Mathematics learning difficulty (MLD) is defined as students' inability to understand mathematical processes and difficulty in performing tasks involving numbers or mathematical symbols (Khin Eng & Sai Hoe, 2021). It is characterized by difficulties in various mathematical skills that cannot be explained by low intelligence, neurological disorders, or inadequate education (Munez et al., 2023; Sudha & Shalini, 2014). Since it is a broad term that covers different types of difficulties that occur in the process of learning mathematics, some researchers classify these difficulties under three categories: a) difficulty using mathematical concepts, which relates the failure to recall technical terms, the inability to express the meaning of terms related to specific concepts, and the difficulty in remembering one or more conditions related to an object, b) difficulty using mathematical principles including mathematical calculations or operations, abstract patterns, and difficulties in explaining the meaning of principles, and c) difficulty solving verbal problems, which involves the knowledge and skills of using concepts and principles (Pramesti & Prasetya, 2021).

The cause of MLD can be recognized when students start to learn mathematics formally in the first grade (Salihu & Räsänen, 2018). The American Psychiatric Association (2013) reported that between 5% and 15% of school children have learning disabilities that prevent them from achieving numerical proficiency. These rates indicate that in a class of 30 students, at least one student shows signs of MLD (Hannell, 2013). Therefore, teachers' awareness of what kind of difficulties students may have in the process of learning mathematics guides how mathematics instruction should be designed (Wijaya et al., 2019). In other words, it is essential for teachers to be knowledgeable about MLD not only to correctly identify students with MLD but also to implement effective practices in the learning environment.

In learning environments, students with MLD, one of the groups of students with special educational needs who cannot fully adapt to standard education programs, need special teaching methods and support services due to the various difficulties they experience in the learning process. The most important tool for meeting the educational needs of students with MLD is Individualized Educational Plan (IEP). IEP is a written document that covers the developmental performance of individuals with special needs, their educational goals and how these goals will be realized and evaluated; it is developed in cooperation with families, teachers and other relevant experts and documents the necessary services and adaptations for the needs of individuals with special needs (Berkant & Atılgan, 2017). The IEP provides an important framework for determining the educational needs of students, setting goals, monitoring progress and adapting the educational program when necessary (Akhanlı et al., 2024). At the same time, it also contains the environment in which the student will receive education, the materials to be used for the student, methods and techniques. Overton (2014) states that while the IEP includes the current academic achievement and functional performance of the student with special needs, this performance includes the effect of the existing inadequacy on the student's participation and achievement in the general education classroom. In other words, the IEP reveals the current performance of students and explains how the inadequacies of students affect their progress in the general education curriculum. Thus, it is possible to determine where the student is in the curriculum and in which areas he/she needs to develop.

Teachers' experiences and perspectives influence their planning and integration of mathematics practices into the learning environment (Cross Francis et al., 2015). Research indicates a relationship between teachers' beliefs and their classroom practices (Beswick, 2017; Sherer et al., 2016). The evidence suggests that structured and controlled classroom practices are necessary to prevent students with MLD from underperforming and to achieve effective learning outcomes. Additionally, the use of intriguing materials can help to capture the attention of these students in mathematics learning environments. For instance, materials that engage multiple senses, such as visual aids like beads or cereals, colored cubes, numbered markers, musical connections, pizza slices, and blocks, are recommended (Kelly, 2020). The use of concrete objects in the teaching process helps students to understand mathematical concepts better. In a meta-analysis, Ise and Schulte-Körne (2013) emphasized that successful outcomes in teaching mathematics to students with MLD can be achieved through individualized instruction, the creation of a structured and hierarchical curriculum, tailoring the lesson to the student, and the use of lots of repetition. However, research indicates that even experienced teachers have limited knowledge about strategies to intervene in MLD (Sousa et al., 2017; Tennant & Tennant, 2010; Wijaya et al., 2019) and lack sufficient awareness about MLD (Chinn & Ashcroft, 2006; Fu & Chin, 2017; Hacısalıhoğlu Karadeniz, 2013).

Students with MLD in learning environments can be considered a significant educational concern. Therefore, it is crucial to identify MLD early and intervene appropriately in advance. Teachers should be knowledgeable about MLD and instructional practices to effectively remediate difficulties and errors. In a learning environment with students of different abilities and learning styles, teachers must be aware of their students' needs and differences. The comprehensive structure of MLD, including various mathematics learning problems, can create knowledge and practice gaps. Most teachers have insufficient knowledge about the characteristics, symptoms, and intervention strategies of MLD (Soo May & Sai Hoe, 2022). Thus, describing teachers' knowledge of MLD and its applications for these students in learning settings is necessary. It is also essential to describe how they design the classroom environment and their preferred methods, techniques, strategies, and materials.

A review of the literature reveals a paucity of research on MLD and a recent increase in research on the topic. A comparison of MLD to dyslexia, another subset of learning difficulties, reveals a significant disparity in the amount of research conducted on each, particularly a 14:1 ratio of studies on dyslexia to studies on MLD (Price & Ansari, 2013). There appears to be a need to research MLD to fill these gaps in the literature. Studies conducted with teachers working with students with MLD in Turkey aimed

to examine the views of teachers and families on the difficulties experienced by students with MLD and the reasons (Temur et al., 2018), the observations and experiences of classroom teachers on MLD (Kaçar, 2018), and the harmony of classroom teachers' opinions on MLD and teaching practices (Avcı, 2020). Some studies have been directed at investigating middle school teachers' perceptions of MLD (Alkan Nurkan & Yazıcı, 2020; Hacısalıhoğlu Karadeniz, 2013). Overall, it was necessary to conduct this study due to the limited number of studies in the literature that addressed knowledge and teaching practices related to MLD. Furthermore, this research aims to contribute to the literature by describing the knowledge and practices of primary school teachers, revealing their knowledge about how teachers define MLD difficulty, the causes of MLD and the effects of MLD on students, evaluating the work teachers do before and during the teaching process, whether they make the environmental organization of the classroom, how they evaluate student achievement, and evaluating how they evaluate student achievement within the framework of classroom practices, and addressing the difficulties experienced in classroom practices in detail. The study will shed light on the knowledge of the teachers about MLD and their practices in the classroom.

METHOD/MATERIALS

Research Design

This research involves a case study that seeks to uncover teachers' knowledge and classroom practices regarding MLD. A case study is a method that involves an in-depth examination of one or more events, environments, programs, social groups, or other interconnected systems (Yıldırım & Şimşek, 2016). Yin (1984) identified four types of case studies: single holistic case designs, single embedded case designs, multiple holistic case designs, and multiple embedded case designs. In single holistic case designs, the unit of analysis can be either a single individual, school, or institution. In a single embedded case design, multiple substrates exist within a single case. On the other hand, the multiple holistic case design is preferred when there is more than one unit of analysis. Similarly, in the multiple embedded case design, there are multiple situations, and each situation is divided into sub-units (Yıldırım & Şimşek, 2016). Teachers' knowledge and classroom practices regarding MLD were considered a single case in the research. Thus, the study is designed as a holistic single case study.

Study Group

The study group for the research was determined using criterion sampling, one of the purposive sampling methods. This method of sampling involves the study of a situation that meets a set of predetermined criteria (Yıldırım & Şimşek, 2016). In the present study, the criteria for working with one or more students with MLD currently in primary school and volunteering to participate in the research. Twenty teachers who fulfilled the criteria were included. To comply with ethical rules, the names of the teachers were kept confidential, and they were identified as T1, T2, and so on. Demographic information about the teachers is given in Table 1.

Table 1. Demographic information about the teachers

		f	%
Gender	Female	6	30
	Male	14	70
Educational level	License	14	70
	Master's degree	6	30
Year of service	6-10 years	1	5
	11-15 years	3	15
	16 years and above	16	80
Branch of teachers	General education teachers	20	100
The classrooms where teachers work	3	10	50
	4	10	50
The socioeconomic status of the school where they work	Middle	13	65
	High	7	35

According to Table 1, 70% of the participants were male, while the remaining 30% were female teachers. Almost a third (30%) of the teachers had a master's degree. 16 of the teachers had 16 or more years of service, while 4 teachers had 1-3 years of service. All teachers (100%) are general education teachers, and two-thirds (65%) work in middle socio-economic status schools. Half of the teachers (%50) work in the third grade and the other half (%50) in the fourth grade.

Data Collection Tools

The researchers developed a semi-structured interview form to collect the research data. The form consisted of two parts. In the first part, the demographic information of the teachers (gender, level of education, and years of service) was obtained; in the second part, open-ended questions were posed to ascertain their knowledge of MLD and their views on classroom practices. The questions in the interview form were presented to the teachers in the same order. Teachers were given questions such as how they define MLD, whether they adjust for students in classroom practices, and if so, what these adjustments may be, and what

kind of practices they use to support mathematical skills. For example: Do you have any information about MLD? How do you define MLD?

Data Collection Process

The interviews were conducted in an available room within the schools where the teachers worked during the appropriate time periods. The interviews were recorded using a voice recorder with the teachers' permission. The interviews typically lasted 45-70 minutes. The interviews were conducted by the second author.

Data Analysis

After the interviews, the teachers' views on the topic were analyzed using the content analysis technique. Content analysis is defined as a systematic and repeatable technique in which some words of a text are combined with smaller content categories through coding based on specific rules (Büyüköztürk et al., 2015). In this study, teachers' responses to each question were first transcribed. Then, words or groups of words were identified following the purpose of the research, and codes were formed; themes were formed by bringing together codes related to each other.

Reliability

To ensure the research's credibility, transferability, reliability, and confirmability, the following measures are taken: Participant verification is necessary to ensure credibility (Maxwell, 2004). To avoid misunderstandings and misinterpretations, the codes, categories, and themes the researcher had created were sent to the teachers for their opinions on whether their comments were correct. Transferability is ensured by providing detailed information about the research (Creswell, 2014). In this regard, how the study group was determined, and the characteristics of the participants were described. Reliability is ensured by more than one researcher coding the data sets and ensuring the consistency between the codes (Yıldırım & Şimşek, 2016). First author listened to the audio recording of all the interviews in order to assess that the transcripts of the interviews were correct. With the decision that the interviews were complete and correct, the research data were analyzed. The teachers' responses to each interview question were examined separately by two researchers. Within the scope of the themes and sub-themes created in the light of the obtained data, the topics with "consensus" and "disagreement" were discussed. If the researchers marked the same category in the relevant question or did not mark any categories related to the question, this was considered a consensus; if the researchers marked a different category from each other, it was considered a disagreement. In this case, the analysis made by the first author is referenced. The reliability calculation of the research was calculated using the formula $\text{reliability} = \text{consensus} / (\text{consensus} + \text{disagreement})$. As a result of the comparison of the coding conducted by the researchers, the percentage of agreement was found to be 0.90. Confirmability was ensured by including direct quotations in the findings section. The formulation of interview questions, data collection, analysis, interpretation of findings, and results were presented. Approval for the ethical compliance of the research was obtained from the university.

FINDINGS

The findings collected from the semi-structured interviews with the teachers are categorized under two main themes. Figure 1 shows the themes and sub-themes.

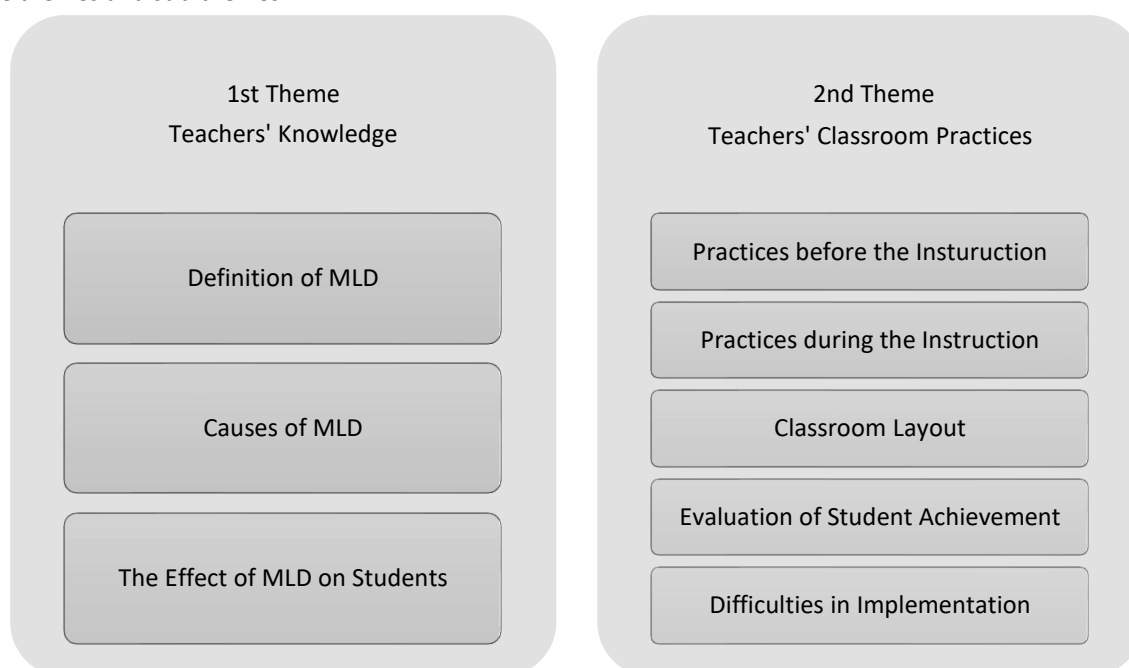


Figure 1. Themes and sub-themes

Definition of MLD

Strikingly, teachers make a distinction between MLD and intellectual disability. Only four teachers emphasized that MLD is related to difficulty learning mathematics despite having average intelligence. T13 said, *"It has difficulty grasping numbers and symbols despite no mental problem. In addition and subtraction, even if the questions are simple, solving them can take significant time."* Additionally, it is often the case that students forget what they have learned. T15 expressed, *"I can define it as a special learning disorder causing difficulty in performing mathematical operations and establishing a relationship between operations although there is no mental problem."*

In particular, two teachers defined MLD as a learning situation that requires repetition. For instance, T6 maintained, *"A student tries to learn subjects later and with more repetition among their peers."* T9 also added *"The situation of learning numbers and operations with numbers for a longer time and with more repetitions than normal."* However, in one of the teachers' definitions, mathematics and reading skills were associated with each other, and it was stated that MLD emerged because of deficiencies in reading skills. T8 expressed, *"MLD can be defined as the inability to visualize the question in the mind or the inability to understand what is read."*

Causes of MLD

Results also revealed that the teachers' knowledge was superficial and general. Notably, the teachers mentioned the reasons for the emergence of other disability groups and frequently mentioned the factors that may cause low achievement while listing the causes of MLD. For example, students' negative attitudes (f=7), teaching difficulties (f=7), intelligence/intellectual disability (f=6), and environment (f=5) were the most underlined concepts. T14 stated, *"Besides mental ability and inability, situations such as learned helplessness and prejudice can be effective. It may also be the inadequacy of materials."* Additionally, T2 said, *"Not using intelligence, when necessary, not making mathematics teaching enjoyable and easy, inexperience of the teacher and leader."* The teachers' explanation of the causes of MLD is shown in Figure 2.



Figure 2. Word cloud explaining the causes of MLD

However, teachers' opinions on the causes of MLD displayed the limitations of their knowledge. Curriculum intensity (f=3), attention deficit (f=3), family indifference (f=3), reading difficulties (f=2), nutrition (f=1), pre-natal and post-natal reasons (f=1), exposure to technology (f=1), socio-economic status of the family (f=1), and hyperactivity (f=1) were among the findings. For example, T7 expressed, *"Negative events before and after birth - falls, accidents, etc. –might affect."* T11: *"S/he may spend too much time with stimuli, such as tablets, computers, phones, etc. Family indifference may also occur."*

The Effects of MLD on Students

All teachers stated that MLD negatively impacts students' education and daily lives. Teachers explained how failing national and international mathematics class exams affects educational life. All the teachers of this opinion underscored the negative manifestation of low mathematics achievement in the classroom on friendships. In this regard, T4 stated, *"It may negatively affect students' educational life in the classroom. They may be exposed to the negative attitudes of their classmates."* Similarly, T10 expressed, *"Mathematics has a major significance in an exam-oriented system. Mathematics questions are pivotal in national and international exams."*

Another notable finding is the teachers' consideration of the psychological effects of students suffering from MLD. Teachers point out that students' failure in mathematics might initiate ridicule and loss of self-confidence. Students with MLD may be subjected to harsh reprimands from teachers, family, and peers and at an increased risk of social isolation. T15 stated, *"Students with MLD can be ridiculed in their educational life. If the teacher is unaware of this issue, negative words can make this distinctive situation of the child more difficult. Likewise, the actions of parents unconscious of the situation may negatively affect the development of the child."* Besides, T16 mentioned, *"It can cause serious psychological problems in children's educational lives. There is a feeling of failure. There is a loss of self-confidence. Since there will inevitably be peer pressure, psychological problems may appear."* One of the teachers implied that the effect on student psychology can only be visible during the school period and will disappear in adulthood. T9 explained, *"The individual with MLD may become disinterested in mathematics over time and gravitate towards verbal fields. However, as they mature, they may become more comfortable with math and use it as needed."*

Another point worth mentioning is the scarcity of teachers (f=2) stating the daily difficulties that students with MLD may experience. It was claimed that mathematics is a crucial skill for daily life, such as telling time, shopping, and figuring out money. For example, T8 said, *"They may have problems while shopping in the canteen and grocery store."* T11 expressed, *"It negatively affects the daily life skills of especially primary school students. They cannot tell the time; they have difficulty in recognizing it. They cannot calculate change, etc."*

Another crucial finding is that teachers have varying views on the impact of MLD on professional life. While some teachers believe that mathematics is essential for every profession, others argue that students may face limitations in their career choices due to MLD. Inferior performance in mathematics can negatively impact career decisions. T17 stated, *"Fields such as medicine, engineering, aerospace, and genetics are commonly considered qualified professions and require a strong background in mathematics. Therefore, students with MLD may face difficulties pursuing a career in these fields."* Additionally, T5 stated, *"Mathematics is essential in every aspect of life, at every level, and for every occupational group. It is as necessary for a painter or a poet as it is for a worker in the industry."* However, some teachers believe that choosing professions without mathematics requirements may mitigate the negative impact of MLD. Therefore, leading students towards different fields could prevent the problem. T2 stated, *"Students who dislike mathematics can compensate for their lack of proficiency by pursuing other fields where they can demonstrate their abilities."* T4 asserted, *"I do not believe that mathematics is necessary for the profession, and therefore, it will not have a negative impact."*

Theme 2: Teachers' Classroom Practices

This section of the paper deals with a thorough explanation of the following sub-themes: Practices before the instruction, practices during the instruction, classroom layout, evaluation of student achievement, and difficulties in implementation under the theme of teachers' classroom practices.

Practices before the Instruction

All the teachers reported that they planned and prepared the lesson before teaching, and they expressed their opinions on planning based on the level of the students, teaching methods, and materials. Teachers followed the curriculum and tailored their lesson plans to students' performance. This situation suggests an overall plan for the whole class without a different plan for students with MLD. Only three teachers indicated that they considered learning outcomes in their preparation. Still, they did not clearly explain whether these learning outcomes were for the whole class or the students with MLD. T1 said, *"We attach importance to photocopy-supported activities in a simple style, following the curriculum and based on our special research."* T18 expressed, *"Some students struggle to sort out the calculations and solve the problems owing to a lack of understanding. We sometimes provide simpler questions for these students."*

Two of the teachers stated that they did not make any preparations. This was because the classes were too crowded, and dealing with the students individually was impossible. T4 expressed, *"I cannot make plans due to the overcrowded class. It is not feasible to provide personalized attention to students."* Likewise, T6 added, *"We cannot plan much."*

Most of the teachers asserted that they prepare their materials before teaching. These are visual materials, presentations, the smartboard, and concrete materials. It is worth mentioning that the teachers considered the materials suitable for the whole class and did not prepare alternative materials. For example, T3 said, *"We support the class with visual materials."* T1 also said, *"We especially get assistance from the smartboard."*

One of the prominent findings might be that one of the teachers decides the method to be used before instruction. However, as with the other teachers' opinions, this teacher's view also addresses the whole class as the target group. To illustrate, T17 expressed: *"I investigate and implement teaching methods. I determine the needs of the students and plan the lesson accordingly."* The findings revealed that teachers did not have sufficient knowledge about lesson plans, materials, and the selection of proper methods-techniques before teaching.

Practices during the Instruction

Teachers shared their opinions about the choice of methods and techniques and the use of materials in the teaching process. Accordingly, lectures, learning by doing, drama, demonstration, play, peer teaching, one-on-one teaching, and question-answer are the methods and techniques that teachers frequently employ. This shows that teachers did not include variety in selecting methods and techniques to support students with MLD. Considering that students with MLD do not learn with the same methods as their peers who show typical development, it is noteworthy that teachers choose methods that are more suitable for the general classroom. Also, teachers mostly gave answers about teaching principles, such as progressing from simple to complex, known to unknown, and concrete to abstract, indicating that the content is gradually presented during instruction. T3 stated, *"I explain the subject from simple to complex, concrete to abstract. I use the learning-by-doing method."* T6 said, *"From simple to difficult, I use one-on-one teaching method."* One of the teachers claimed that diversity may not be incorporated in selecting method techniques to support students with MLD. T4 explained, *"Technically, I involve these students in the techniques I apply in class. So, they are not isolated from the class."*

Most teachers utilize the materials available in the classroom during the teaching process. They prefer concrete materials to enhance understanding and encourage student interaction. For instance, T7 stated, *"I use math boxes for addition, toy houses for ones and tens, walnuts, pencils, and fraction sets for division. In problem-solving, it is beneficial to explain by using the shapes."* T19 added that *"I use many materials in the lesson. I bring the student together with the materials."*

The findings also stress that some teachers selected materials based on the student's level. T11 expressed, *"I first prepare an IEP for my student in cooperation with the counselor. I assess their mathematical performance and teach basic math using the materials available depending on the data."* T5 stated, *"I prepare materials appropriate for the student's level."* The statement underlines that only one of the teachers prepares materials.

A teacher mentioned that visualization in teaching is essential and can be achieved using materials such as concept maps, graphs, and figures to facilitate student understanding. T15 explained, *"Using concept maps, graphs, and figures can facilitate the participation of students with special needs in the learning process."* A teacher also stressed technology-supported teaching, stating, *"I utilize presentations, slides, and videos."*

Classroom Layout

Most teachers reported making physical classroom arrangements, such as colorful boards, information signs, and posters, to summarize or make the subject more concrete. For instance, T14 stated, *"I hang big pictures summarizing the subject on the board."* Similarly, T2 said, *"I utilize boards, posters, and colorful signs to convey information."*

However, teachers use visual materials and decorate the classroom with mathematics activities related to the subject to appeal to multiple senses. For instance, they use colors and shapes to draw students' attention to the subject. T20 stated, *"I decorate the classroom with mathematics activities that appeal to more sensory organs. Visual aids are an effective tool for concretizing abstract mathematical concepts. Such aids are beneficial for students who struggle with addition."* T7 expressed, *"I present activities related to numbers, fractions, rhythmic counting, or geometric shapes on the classroom walls."* Cubes, colored papers, orange slicing, graphs, shapes, concept maps, and number tables are also mentioned by the teachers. T15 added, *"I primarily use visual concept maps, graphics, and shapes in activities."* T10 said, *"I use cubes for addition and real-life examples, such as slicing an orange in the classroom, to teach fractions."*

Despite being the least frequent in teachers' evaluations, the most notable activities were reading books and creating a classroom corner. T17 emphasized the role of reading books in concretizing the subject, stating, *"I read books with mathematical context; I tell stories. Thus, I concretize the subjects."* T9 discovered that learning corners facilitated the learning process and increased student engagement. T9 stated, *"I prepare rhythmic counting corners to increase the comprehensibility of the subject."*

Another less-frequently mentioned aspect is the seating position of the students in the classroom. Arrangements for students to sit in the front rows are necessary to ensure careful listening and reduce sight-hearing problems. T19 said, *"I seat students with LD in the front rows so that they can see the board better and listen more carefully."* T8 expressed, *"As classroom layout can effectively address sight and hearing problems, I provide solutions for these issues."*

The findings indicate that four teachers did not make physical arrangements in the classroom. The reasons were crowded classrooms, yearly changes, and excessive costs. For example, T5, *"There are limited opportunities for such arrangements due to the necessary costs and the need for permission from the school administration. It can be challenging to arrange as the classes*

change every year." T6 said, *"I cannot say that I made arrangements since the classrooms are overcrowded."* In general, the physical arrangements made by teachers to support students with MLD appear quite limited.

Evaluation of Student Achievement

Teachers consider the student's level when assessing their achievements. They use level-appropriate questions, worksheets, alternative question types, short answer questions with simple operations, quizzes, and tests to evaluate success. For instance, T8 stated, *"In my evaluation, I assess if students can solve questions of their levels as well as they solve questions on the board."* T13 also explained, *"My assessment criteria are based on student's abilities, measuring at their level. I design exams or tests appropriate to their level."*

Teachers use observation as another method to evaluate student achievement. Observation is used to assess the student's performance in the classroom and track their progress. According to T16, *"In our class, evaluations are based on observation. I follow the development in the process."* T4, *"I observe their performance in the classroom."*

One remarkable finding regarding teachers' evaluation is that only two teachers reported considering IEP outcomes. These teachers stated that they do not evaluate students on class outcomes but on whether they have achieved the set goal. T11 explained, *"I evaluate the achievements of these students on the IEP objectives I have prepared rather than comparing them to the whole class or MoNE objectives."* T17 expressed, *"I assess whether the student has met the target set for them, such as successfully performing addition with numbers from one to twenty."*

When assessing the performance of students with MLD, some teachers made no adjustments. Methods used to assess the whole class include written tests, activity sheets, class participation, scales, and oral questioning and answering. T5 stated, *"I evaluate the situation in the class and the written exams."* Besides, T14 said, *"I assess the written or oral question and answer and participation in-class activities."* The findings indicate that teachers have limitations in assessing students' achievement with MLD, as evidenced by their practice.

Difficulties in Implementation

All teachers encountered difficulties when teaching mathematics to students with MLD. Student-related concerns construed these difficulties. For example, the factors hindering teachers from performing effective instruction include students below grade level, distraction, forgetfulness, indifference to the lesson and homework, prejudice, negative attitudes, learned helplessness, and lack of self-confidence. Teaching students below the grade level was considered the biggest challenge of the abovementioned factors. Accordingly, T5 stated, *"Teachers have difficulty completing coursework and giving adequate attention to students who need extra hours of instruction. Falling behind causes greater problems as the gap expands over time."* T3 also added, *"As mathematics is an abstract field, students often encounter difficulties in learning. This can hinder their ability to progress to new topics."*

Another finding reveals that teachers are challenged by the lack of classroom materials. This issue is consistent with the sub-theme of teaching processes. The unavailability of materials may negatively impact the teaching process. T1 stated, *"Sometimes we may have problems in terms of materials. Mathematics classes should be offered in every school, and special conditions and opportunities should be provided to develop this skill."* T11 expressed, *"We have difficulty accessing sufficient visuals or materials for each subject."* This suggests that teachers use pre-made materials rather than creating materials tailored to the needs of students with MLD.

Furthermore, teachers emphasize the teacher-related reasons for the difficulties in practice. Teachers considered their knowledge and skills about MLD inadequate. For example, T10 said, *"We have difficulties because of inadequate knowledge and skills about this subject. Moreover, teachers found it problematic to prepare questions and activities, assignments, and assessments for students with MLD."* T16 stated, *"Preparing activities and questions for them is difficult and demanding."* T2 explained, *"It is hard to teach, evaluate, and assign homework at their level."* One of the teachers emphasized the limitations of teaching techniques. T17, *"The lack of appropriate teaching techniques for these students often makes it difficult for us to plan and manage the process."*

Another issue that teachers encountered was related to the learning environment. Reasons cited for difficulties included missed deadlines, classroom management, and overcrowded classrooms. Teachers linked the problems with the low level of the students most. Correspondingly, T4 stated, *"The fact that they are below the class level causes difficulties in classroom management. In addition, the overcrowded classes make it difficult for us to communicate with each student and prevent communication with students with MLD. Consequently, less communication means less success."* T6 mentioned *"There needs to be repetition and question-solving. We have time problems in covering the subjects."* In addition, the teachers also highlighted the reasons stemming from the families. The family's lack of interest and refusal to accept support negatively affected the practices. For example, T18 said, *"Sometimes the student does not make the necessary effort to learn and does not stay for one-on-one instruction. Sometimes, the family does not allow it. They should get assistance from an expert, but they do not accept it."* T7 said, *"If the family lacks patience and help for the child, the teacher cannot assist at school."*

DISCUSSION, CONCLUSION AND SUGGESTIONS

Two main themes were identified to determine teachers' knowledge about MLD and their classroom practices: teachers' knowledge and teachers' classroom practices.

Teachers' Knowledge

Teachers' knowledge was analyzed under the sub-themes of the definition of MLD, causes of MLD, and effects of MLD on students. It was found that teachers associated the concept of MLD with deficiencies in mathematical skills, numbers, problem-solving, four operations, calculations, and difficulty in learning mathematical concepts. Only four teachers emphasized "the absence of mental deficiency" in the definition. Thus, teachers can be considered to have conceptual knowledge limitations. An analysis of teachers' definitions of MLD revealed that they mainly mentioned the commonly failed skills and the symptoms of MLD instead of using distinctive and specific expressions. In the literature, some studies unearthed that teachers define MLD in an equivalent way (Avcı, 2020; Hacısalihoğlu Karadeniz, 2013; Sezer & Akin, 2011).

In the DSM-V, learning disability is defined as having average or above-normal intelligence and not caused by any disability or environmental conditions. IDEA (2004) emphasizes that learning disabilities do not include learning problems caused by intellectual disability, emotional disturbance, or environmental, cultural, or economic disadvantage. Thus, MLD refers to underachievement despite normal and abnormal intelligence, emotional stability, educational opportunities, and motivation. The study revealed that teachers explained the causes of MLD in terms of factors such as intelligence, intellectual disability, genetics, environment, curriculum intensity, exposure to technology, lack of instruction, lack of attention, and nutrition. Their knowledge of the subject was relatively superficial and inadequate. The teachers' statements displayed that they attributed MLD to other disability groups and numerous factors that could cause low achievement. Teachers' lack of knowledge about the causes of MLD is often emphasized (Kaçar, 2018). Teachers' misconceptions about learning disabilities and difficulties may also initiate this situation.

Another point is that teachers have limited knowledge about how MLD impacts students' educational, daily, and professional lives. It was highlighted that MLD leads to low academic performance and deteriorating relationships with teachers, family, and peers. The teachers' statements about the impact of academic failure on students' psychology (e.g., loss of self-confidence, isolation, pressure) demonstrated that they were conscious of the problem. However, teachers' statements about the impact of MLD on everyday life were quite limited. Mathematics is needed in many areas, such as calculating distances and costs, scheming dates and calendars, and using numerical expressions in our daily needs. It was ascertained that the teachers could superficially explain the prominence of mathematics, stating that it is only necessary to tell the time, shop, and calculate changes. The greatest confusion among teachers is related to the impact of MLD on professional life. The significance of math skills in students' lives is evident due to their impact on school attendance and their requirements in various occupational groups. For example, Heine et al. (2013) identified a link between poor mathematical skills and psychosocial and economic risks, indicating that those affected might drop out of school, and only a minority are employed full-time at age 30. In this research context, it can be argued that teachers are not fully aware of MLD's impact on career choices.

Teachers' Classroom Practices

Teachers' classroom practices were examined under the sub-themes of practices before teaching, practices during the teaching process, classroom layout, evaluation of student achievement, and difficulties in implementation.

Consequently, it was found that the teachers do not make in-class applications for students with MLD. This suggests that the teachers created a plan without considering the performance of the students with MLD in the classroom practice. Remarkably, some teachers did not make any adjustments to their lesson planning. The teachers' preparation of materials and teaching methods also displayed an analogous situation. It was noted that teachers planned their lessons through visuals, smart boards, and concrete materials in the classroom environment. However, none of the teachers prepared materials to address the needs and content of students with MLD. The choice and use of a single teaching method for the whole class suggests that the teachers' preparation process before teaching is inadequate. In a study conducted by Özkubat et al. (2021) with teachers working with students with learning difficulties, it was revealed that teachers planned the lesson in the same way for students with and without learning difficulties. One of the important results that emerged within the scope of the research is that the methods used by teachers may have applications that they use not only for students with MLD, but also for all students. Such applications also contribute to students with MLD in the classroom environment. It can be thought that teachers who adopt the Universal Learning Design in their classroom applications by taking into account students with different abilities or different learning characteristics. However, when the results obtained from the research are examined from a detailed framework, it is seen that teachers do not make any special applications for students with MLD.

The research yielded how teachers conducted the teaching process, including lectures, learning by doing, drama, demonstration, play, peer teaching, one-on-one teaching, and question-answer method. As found in similar studies, teachers frequently use these methods (e.g., Kaçar, 2018; Özkubat et al., 2021; Temur et al., 2018). It is noted that the methods and techniques teachers use are quite limited and inappropriate. This result is like those in the literature (Avcı, 2020; Kaçar, 2018). Research conducted on students with MLD shows that these students cannot learn with the same methods and at the same speed as their peers with normal development (Re et al., 2014). It is inevitable that the instructional strategies used to support the individual mathematics learning of students with learning difficulties will vary (Akçin, 2019). For example, it is recommended to

focus on instructional strategies such as explicit teaching, strategy teaching, direct teaching, metacognitive approach and supportive teaching (Berkeley et al., 2010, as well as practices such as direct teaching, modeling, providing guided application opportunities, adjusting progress, error correction and recording progress (Akçin, 2019). The study also revealed that the question-answer method was the most preferred by the teachers. The use of simple questions adapted to the level of the students indicated that some teachers made adaptations, yet only partially. However, the learning characteristics of students with MLD are often neglected as most teachers teach to the whole class. These findings are supported by other studies in the literature (Acar & Hiğde, 2018). The second issue addressed in the teaching process concerns the teaching materials. According to the study, teachers reported using pre-made materials that offer visual aids and address multiple senses. However, the study also revealed that teachers lack planning and adapting materials. Some teachers did mention adapting the materials according to the performance and level of the students.

In terms of classroom layout, teachers are found to decorate the classroom with mathematical activities appropriate to the subject and use colorful bulletin boards, and posters. They also stated that these arrangements were necessary for engaging multiple senses and aiding visualization. Book reading and rhythmic counting corners were less favored in the classrooms. Only two teachers arranged for students to sit in the front rows. Previous studies have mentioned that students were seated close to the teacher, and physical adaptations were made in the classroom (e.g., Koç, 2018; Özkubat et al., 2021). A significant finding of this study is that some teachers did not adjust the physical arrangement of the classroom. The teachers listed several reasons for not adjusting, including the need to save money, changing the classroom every year, and overcrowded classrooms.

The research also found limitations in teachers' ability to assess student achievement. The preferred assessment methods include level-appropriate questioning, worksheets, alternate questioning, short answer questions with simple operations, quizzes, and tests. Only two teachers stated that they consider the results of the IEP and do not evaluate students solely on class results but also on whether the student has reached the set objective. Teachers use various methods to evaluate student success, including whole-class written exams, activity sheets, class participation, scales, and oral question and answer. However, these methods lack regulation and consistency in assessing student achievement. The results of the study revealed shortcomings and limitations in teachers' classroom practices for students with MLD, consistent with previous studies (Acar & Hiğde, 2018; Avci, 2020; Sezer & Akin, 2011).

The research found that teachers experienced problems in their practice due to student-related reasons, teachers, learning environment, material insufficiency, and families. It is noteworthy that teachers emphasize student issues, following previous studies (Hacısalıhoğlu Karadeniz, 2013; Sezer & Akin, 2011). Teachers' preference for materials can be inferred from their reports of difficulties due to lack of materials. These difficulties include keeping up with the pace, managing the classroom, and dealing with overcrowded classes caused by students falling back and needing too much repetition. Additionally, the research indicates that the absence of family support and indifference toward the student is a result that aligns with the literature (e.g., Avci, 2020).

As a result of the study, it was determined that teachers had insufficient knowledge about MLD and their classroom practices were also inadequate. Therefore, seminars and trainings can be organized to increase teachers' knowledge about MLD. In addition, in order to overcome teachers' limitations in classroom practices, in-service courses on methods and techniques that can be used in mathematics lessons and on developing materials that appeal to multiple senses can be planned by the Ministry of National Education. By planning these courses regularly, teachers' theoretical and practical knowledge can be increased. It is also important to encourage teachers to attend seminars and in-service trainings on MLD. Guidance booklets can be prepared for teachers to be informed about MLD and the methods, techniques, strategies and materials they can use in the classroom.

The research focused on the determination of teachers' knowledge about MLD and their views on classroom practices. It may be recommended that future research should be more comprehensive and use different data collection tools that provide detailed and rich information on the subject. In addition, research can be conducted to examine the knowledge and classroom practices of subject teachers at the middle and high school levels.

Declaration of Conflicting Interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The authors received no financial support for the research, author-ship, and/or publication of this article.

Statements of publication ethics

We hereby declare that the study has not unethical issues and that research and publication ethics have been observed carefully.

Researchers' contribution rate

The study was conducted and reported with equal collaboration of the researchers.

Ethics Committee Approval Information

Within the scope of ethics committee permission; Tokat Gaziosmanpaşa University Social and Human Science Research Ethics Committee, 12/03/2024, 05.28 numbered document was obtained.

REFERENCES

- Acar, E., & Hiğde, A. (2018). The examination of an elementary school student having dyscalculia in classroom environment. *Western Anatolia Journal of Educational Sciences*, 9(2), 102-119.
- Akçin, N. (2019). Öğrenme güçlüğü olan öğrenciler için genel eğitim sınıflarında uygulanacak etkili öğretimsel stratejiler ve öğretimsel düzenlemeler [Effective instructional strategies and instructional arrangements for students with learning disabilities in general education classrooms]. In E. R. Özmen & A. Ataman (Eds.), *Öğrenme güçlüğü ve özel yetenek* [Learning disability and special ability] (pp. 55-89). Maya Akademi.
- Akhanlı, M., Tama, G., & Bülbül, M. (2024). Özel gereksinimli öğrencilerde bireyselleştirilmiş eğitim. *International Journal of Social and Humanities Sciences Research (JSHSR)*, 11(107), 979-988. <https://doi.org/10.5281/zenodo.11405475>
- Alkan Nurkan, M., & Yazıcı, F. (2020). A case study on determination of the mathematics teachers' awareness of mathematics learning disability (Dyscalculia). *Journal of Contemporary Administrative Science*, 7(1), 95-109.
- American Psychiatric Association. (2013). *Diagnostic and statistical DSM-5*. Washington, DC.
- Avcı, A. (2020). *Evaluation of teaching applications for primary school students living mathematics learning disability* [Master thesis]. Dokuz Eylül University.
- Berkant, H., G. & Atılın, G. (2017). Sınıf öğretmenlerinin kaynaştırma eğitimine yönelik yaşadıkları sorunlar ve çözüm önerileri [Problems and solutions proposed by classroom teachers for inclusive education]. *Journal of Educational Reflections*, 1(1), 13-25.
- Berkeley, S., Scruggs, T. E., & Mastropieri M. A. (2010). Reading comprehension instruction for students with learning disabilities, 1995-2006: A meta-analysis. *Remedial and Special Education*, 31(6), 423-436. <https://doi.org/10.1177/0741932509355988>
- Beswick, K. (2017). Raising attainment: What might we learn from teachers' beliefs about their best and worst mathematics students? In C. Andra, D. Brunetto, E. Levenson, & P. Liljedahl (Eds.), *Teaching and learning in mathematics classrooms* (pp. 95-106). Springer International Publishing.
- Büyüköztürk, Ş., Kılıç Çakmak, E., Akgün, Ö., Karadeniz, Ş., & Demirel, F. (2015). *Bilimsel araştırma yöntemleri* [Scientific research methods]. Pegem Akademi Publishing.
- Chinn, S., & Ashcroft, R. (2006). *Mathematics for dyslexics: Including dyscalculia*. Wiley
- Creswell, J. W. (2014). *Research design: Qualitative, quantitative, and mixed methods approaches*. SAGE Publications.
- Cross Francis, D., Rapacki, L., & Eker, A. (2015). The individual, the context, and practice: A review of the research on teachers' beliefs related to mathematics. In H. Fives, & M. Gregorie Gill (Eds.), *International handbook of research on teachers' beliefs* (pp. 336-352). Routledge.
- Fu, S. H., & Chin, K. E. (2017). An online survey research regarding awareness of dyscalculia among educators in Sandakan district. *Int. J. Acad. Res. Progress. Educ. Dev.* 6(2), 1-10.
- Hacısalıhoğlu Karadeniz, M. (2013). Assessment of teachers' views on students with the disorder of "Dyscalculia". *E-Journal of New World Sciences Academy Social Sciences*, 8(2), 193-208.
- Hannell, G. (2013). *Dyscalculia: Action plans for successful learning in mathematics*. Routledge.
- Heine, A., Wissmann, J., Tamm, S., et al. (2013). An electrophysiological investigation of non-symbolic magnitude processing: Numerical distance effects in children with and without mathematical learning disabilities. *Cortex*, 49, 2162-77.
- Individuals with Disabilities Education Improvement Act of 2004 (IDEA) (2004). Pub. L. No. 108-446. Retrieved from <http://www.flspedlaw.com/idea.2004.all.pdf>.
- Ise, E., & Schulte-Körne, G. (2013). Symptomatik, diagnostik und behandlung der rechenstörung. *Zeitschrift für Kinder- und Jugendpsychiatrie und Psychotherapie*, 41(4), 271-282.
- Kaçar, H. (2018). *Analysis of primary school students' mathematics learning disabilities according to their teachers' observations and experiences* [Master thesis]. Necmettin Erbakan University.
- Kelly, K. (2020). *Identifying, assessing, and supporting learners with dyscalculia*. SAGE.
- Khin Eng, C., & Sai Hoe, F. (2021). Exploring the implementation of an intervention for a pupil with mathematical learning difficulties: A case study. *Journal on Mathematics Education*, 12(3), 531-546.
- Koç, B. (2018). *An action research on teaching addition and subtraction to students with dyscalculia* [Doctoral Thesis]. Necmettin Erbakan University.
- Maxwell, J. A. (2004). Using qualitative methods for casual explanation. *Field Methods*, 16(3), 243-264.
- Munez, D., Bull, R., Lee, K., & Ruiz, C. (2023). Heterogeneity in children at risk of math learning difficulties. *Child Development*, 94, 1033-1048.
- Overton, T. (2014). *Assessing learners with special needs an applied approach*. Pearson Education.

- Özkubat, U., Sanır, H., & Özmen, E. R. (2021). Teachers' opinions on instructional adaptations for students with learning disabilities. *Ahi Evran Üniversitesi Sosyal Bilimler Enstitüsü Dergisi*, 7(3), 881-900.
- Pramesti, C., & Prasetya, A. (2021). Analisis tingkat kesulitan belajar matematika siswa dalam menggunakan prinsip matematis. *Edumatica : Jurnal Pendidikan Matematika*, 11(02), 9–17.
- Price, G. R., & Ansari, D. (2013). Dyscalculia: Characteristics, causes, and treatments. *Numeracy*, 6(1), 1–16.
- Re, A. M., Pedron, M., Tressoldi, P. E., & Lucangeli, D. (2014). Response to specific training for students with different levels of mathematical difficulties. *Exceptional Children*, 80(3), 337-352. <https://doi.org/10.1177/0014402914522424>
- Salihu, L., & Räsänen, P. (2018). Mathematics skills of kosovar primary school children: A special view on children with mathematical learning difficulties. *İjee*, 10(4), 421-430.
- Scherer, P., Beswick, K., DeBlois, L., Healy, L., & Opitz, E. M. (2016). Assistance of students with mathematical learning difficulties: How can research support practice? *ZDM Mathematics Education*, 48(5), 633-649.
- Sezer, S., & Akin, A. (2011). Teachers' opinions about dyscalculia seen in the students between the ages of 6-14. *Elementary Education Online*, 10(2), 757-775.
- Soo May, Y., & Sai Hoe, F. (2022). Teachers' efficacy to teach students with mathematics learning difficulties. *Journal of Mathematics Instruction, Social Research and Opinion*, 1(3), 117-130.
- Sousa, P., Dias, P. C., & Cadime, I. (2017). Predictors of primary school teachers' knowledge about developmental dyscalculia. *Eur. J. Spec. Needs Educ.*, 32(2), 204–220.
- Sudha, P., & Shalini, A. (2014). Dyscalculia: A specific learning disability among children. *International Journal of Advanced Scientific and Technical Research*, 4(2), 912-918.
- Temur, O. D., Turgut, S., Ozdemir, K. (2018). Teachers and parents' perception about learning difficulties in mathematics: A case study. *International Online Journal of Educational Sciences*, 10(4), 126-148.
- Tennant, L.J., & Tennant, R.F. (2010). Dyscalculia: More than mathematics phobia. *Middle East Educator*, 14, 46–49.
- Wijaya, A., Retnawati, H., Setyaningrum, W., Aoyama, K., & Sugiman. (2019). Diagnosing students' learning difficulties in the eyes of Indonesian mathematics teachers. *J. Math. Educ.*, 10(3), 357–364.
- Yıldırım, A., & Şimşek, H. (2016). *Sosyal bilimlerde nitel araştırma yöntemleri* [Qualitative research methods in social sciences]. Seçkin Publishing.
- Yin, R. (1984). *Case study research: Design and methods*. Sage Publications.



| Research Article / Araştırma Makalesi |

Exploring Preservice Science Teachers' Socioscientific Reasoning Competence

Fen Bilimleri Öğretmen Adaylarının Sosyobilimsel Muhakeme Becerileri

Süleyman ÇİTE¹, Ahmet ÜNAL², Selçuk ŞAHİNGÖZ³

Keywords

1. Socioscientific issues
2. Preservice Science Teachers
3. Socioscientific reasoning

Anahtar Kelimeler

1. Sosyobilimsel konular
2. Fen bilimleri öğretmen adayları
3. Sosyobilimsel muhakeme becerileri

Received/Başvuru Tarihi

10.03.2025

Accepted / Kabul Tarihi

10.10.2025

Abstract

Purpose: Contemporary science education standards underscore the importance of scientific literacy. To this end, science education researchers have been discussing what scientific literacy entails in a globalized world. A recent approach defines scientific literacy not only as the ability to comprehend and apply science concepts but also as the competence to construct arguments about complex, science-based societal issues. This requires scientifically literate individuals to go beyond science concepts and examine societal issues from multiple perspectives, such as economic, social, and ethical, a practice known as socioscientific reasoning. Undoubtedly, developing students' socioscientific reasoning depends on having science teachers who are themselves competent socioscientific reasoners. In this respect, this study aims to document the socioscientific reasoning competencies of preservice science teachers through open-ended tasks incorporating societal issues.

Design/Methodology/Approach: To assess preservice science teachers' socioscientific reasoning competencies, an open-ended response task with two different socioscientific scenarios was designed and administered to 76 third and fourth year teacher candidates. Responses were analyzed with using inductive analysis method. Through constant comparison method, the responses of students were categorized into three level competency.

Findings: The results show that while preservice teachers understood the complexity of these societal issues, their reasoning was naive in approaching the issues from the perspectives of different stakeholders, appreciating the need for further inquiry, and skeptically examining the issue.

Highlights: The study concludes with implications for preservice science teacher education programs.

Öz

Çalışmanın amacı: Çağdaş fen öğretim standartları, bilimsel okuryazarlığın önemini vurgulamaktadır. Fen eğitimi araştırmacıları, küreselleşen bir dünyada bilimsel okuryazarlığın ne anlama geldiğini tartışmaktadır. Yakın tarihli bir yaklaşım, bilimsel okuryazarlığı yalnızca bilimsel kavramları anlama ve uygulama becerisi olarak değil, aynı zamanda karmaşık, bilim temelli toplumsal konular hakkında argümanlar oluşturma yetkinliği olarak tanımlar. Bu, bilimsel olarak okuryazar bireylerin yalnızca bilimsel kavramlarla sınırlı kalmayıp; ekonomik, sosyal ve etik gibi farklı açılardan toplumsal sorunları incelemelerini gerektirir ki bu uygulama "sosyobilimsel muhakeme" olarak adlandırılır. Hiç kuşkusuz, öğrencilerin sosyobilimsel muhakeme becerilerini geliştirebilmeleri, bu konuda yetkin fen öğretmenlerinin varlığına bağlıdır. Bu bağlamda, bu çalışma fen bilimleri öğretmen adaylarının sosyobilimsel muhakeme yeterliklerini sosyobilimsel sorunları içeren açık uçlu sorular aracılığıyla ortaya koymayı amaçlamaktadır.

Materyal ve Yöntem: Fen bilimleri öğretmen adaylarının sosyobilimsel muhakeme yeterliklerini değerlendirmek amacıyla, iki farklı sosyobilimsel senaryoyu içeren açık uçlu bir ölçme aracı tasarlanmış ve 3. ve 4. sınıfta öğrenim gören 76 öğretmen adayına uygulanmıştır. Yanıtlar, tümevarımsal analiz yöntemi kullanılarak analiz edilmiştir. Sürekli karşılaştırma yöntemi aracılığıyla, öğrencilerin yanıtları üç yeterlik düzeyine kategorize edilmiştir.

Bulgular: Sonuçlar, öğretmen adaylarının sosyobilimsel konuların karmaşıklığını anladıklarını, ancak farklı paydaşların bakış açılarından soruna yaklaşma, daha fazla araştırmaya duyulan ihtiyacı takdir etme ve konuyu sorgulayıcı bir şekilde inceleme konularında muhakemelerinin oldukça yüzeysel olduğunu göstermektedir.

Önemli Vurgular: Çalışma, fen öğretmeni yetiştirme programları için çıkarımlarla (önerilerle) sona ermektedir.

¹ Corresponded Author, Kastamonu University, Faculty of Education, Department of Mathematics and Science Education, Kastamonu, Türkiye; <https://orcid.org/0000-0002-1359-3621>

² Kastamonu University, Faculty of Education, Department of Elementary Education, Kastamonu, Türkiye; <https://orcid.org/0000-0001-8617-6602>

³ Gazi University, Gazi Faculty of Education, Department of Mathematics and Science Education, Ankara, Türkiye; <https://orcid.org/0000-0003-4884-7588>

INTRODUCTION

Recent reform efforts in science education have emphasized the importance of developing scientifically literate citizens (Ministry of National Education [MoNE], 2024; National Research Council [NRC], 2012). Although there are variations in the definition of what a scientifically literate person can do, the consensus definition describes it as the ability to construct arguments regarding societal issues. These issues, as it called socioscientific issues (SSI) involve not only science but also social, ethical, economic, and cultural aspects that must be considered to make a well-informed decision (Roberts, 2007; Roberts & Bybee, 2014; Sadler, 2004). Contrary to well-known definitions of scientific literacy, which often involve only the development of conceptual understanding and science practices, contemporary perspectives focus on an individual's ability to be a competent socioscientific reasoner when addressing societal issues that involve science (Zeidler & Sadler, 2023).

Today societies face global as well as local challenges like pandemics, climate change, wildfires or food scarcity, which are all considered as SSI. While science might have solutions to these issues, for the proposed solutions to these problems to become effective, it is essential that other dimensions of the issue—such as economic, social, moral, and ethical aspects—are also taken into account within the solutions. Otherwise, even if the proposed solutions have the potential to be scientifically effective, they will fail to produce the expected impact if they are not socially accepted (Zeidler & Sadler, 2023). Therefore, in order to address both the current problems we are facing whose effects are still ongoing and the potential problems we may encounter in the future in an effective manner at the societal level, the development of scientifically literate individuals should be regarded as a major objective within science education programs (Sadler & Zeidler, 2009).

The need to investigate the cognitive processes that a scientifically literate individual should employ when reasoning about SSI, as well as to examine how their reasoning develops following an instructional process focused on SSI has become an important issue to investigate. In response to this challenge, many science educators have investigated students' gains in critical thinking, informal reasoning, decision-making, and argumentation as a result of engagement in instruction designed around socioscientific issues (Zohar & Nemet, 2002; Sadler & Zeidler, 2005; Fowler, Zeidler, & Sadler, 2009; Zeidler, Herman, Ruzek, Linder & Lin, 2013). While these higher-order learning outcomes are valuable for students to develop, they do not fully reflect the competencies or cognitive processes necessary for crafting informed solutions to SSI (Romine, Sadler, & Kinslow, 2017). In response to this, Sadler and colleagues (2007) conceptualized socioscientific reasoning (SSR) as a theoretical construct that defines the thinking processes necessary for navigating a socioscientific issue, comprising four dimensions: appreciating the complexity of the issue, examining it from multiple perspectives, recognizing the need for further inquiry, and maintaining a critical stance toward the issue.

Given that teachers have a central role in science classrooms, they themselves need to be competent socioscientific reasoners to develop learners' ability to form convincing arguments about societal issues. Understanding the current state of preservice science teachers' SSR provide suggestions for improvement in both preservice science teachers' education curriculum as well as professional development opportunities for science teachers in general. For this reason, this study aims to explore preservice science teachers' SSR through open ended scenario based assessments.

Conceptual Framework

Many studies have examined the development of students' reasoning as a result of engagement in SSI-based instruction. The types of reasoning examined in these studies have included informal reasoning (Sadler & Zeidler, 2005), argumentation (Zohar & Nemet, 2002), decision making (Patronis, Potari & Spiliotopoulou, 1999), and critical thinking (Kolstø, Paulsen & Mestad, 2024). However, Sadler and colleagues (2007) have argued that the type of reasoning required for advanced engagement with SSI to form informed decisions about such issues involves different cognitive as well as affective processes, which they call socioscientific reasoning (SSR). In this study, we adopted this conceptualization of SSR to document preservice science teachers' reasoning abilities when faced with SSI.

Socioscientific Reasoning (SSR)

According to Sadler and colleagues (2007), socioscientific reasoning (SSR) encompasses four key dimensions: acknowledging the complexity of issues, considering issues from multiple perspectives, recognizing that issues require continuous inquiry, and maintaining skepticism toward the information presented.

By nature, SSI are characterized as complex and contentious issues that are open-ended and cannot be resolved through simple, clear-cut solutions (Sadler, 2004). An advanced form of reasoning that reflects an appreciation of the complexity of SSI involves avoiding a single, linear cause-and-effect approach to these issues. Instead, it requires recognizing the multiple dynamic interactions inherent in the issue and formulating solutions that take into account the contextual details of the SSI as well (Sadler, Barab & Scott, 2007). Advancement within the complexity dimension of SSR is marked by a transition from basic causal reasoning to reflective analysis, where learners consider and weigh complex, and frequently conflicting, scientific and social information surrounding an SSI (Kinslow, Sadler & Nguyen, 2019).

The multiple perspectives aspect of SSR requires examining an issue from the viewpoints of the different stakeholders involved in the SSI (Zeidler & Sadler, 2010). One of the defining features of SSI is that stakeholders hold diverse interests, beliefs, ideas, and biases regarding the resolution of such issues. Consequently, advanced reasoning about SSI involves understanding these varying interests, beliefs, ideas, and biases and formulating solutions that take these challenges into account. In the context of perspective-taking within SSR, more advanced engagement is reflected in the ability to evaluate issues and possible solutions from varied perspectives and to critically consider counterarguments to one's stance; by comparison, less advanced engagement is marked by a reliance on prompts to recognize differing views and difficulty in moving beyond a personal frame of reference (Sadler, et. al., 2007).

The inquiry aspect of SSR highlights that SSI are complex issues that scientific as well as social dimensions of them require ongoing exploration (Sadler, et. al., 2007). Consequently, there always are unanswered questions about scientific as well as social aspects of SSI and being able to aware of and recognize such unanswered aspects of the issue being explored is an another competence within SSR. Competent learners on the inquiry aspect of SSR are expected to formulate targeted questions or requests for information about uncertain parts of SSI and to develop a plan for addressing at least some of these uncertainties. Ideally, this inquiry should encompass both scientific and social aspects of issue. By contrast, students demonstrating a naive approach to inquiry might merely identify or list the uncertain aspects of an issue without going beyond that (Kinslow, et. al., 2019).

Skepticism constitutes another key dimension of SSR, requiring awareness of potential biases in the information provided by various stakeholders engaged in the issue (Sadler, et. al., 2007). For instance, students should analyze media sources with a critical lens, identify possible stakeholder biases, and question the soundness of conclusions or claims related to scientific dimension of SSI (Kinslow, et. al., 2019). Competent students in this aspect of SSR should be able to recognize potential biases and their sources, and propose ways to address these challenges in order to make well-informed decisions about SSI.

Following Sadler and colleagues' (2007) conceptualization of the SSR construct, a growing body of research has investigated the SSR competencies of diverse learner groups, ranging from middle school students to preservice and in-service teachers, and has explored how these competencies develop through participation in SSI-based instruction. SSR was first employed to evaluate the reasoning of middle school students while a virtual exploration of an environmental SSI (Sadler, et. al., 2007). This pioneer study helped researchers to examine how students applied socioscientific reasoning skills after participating in a ten-day SSI instructional unit. The findings indicated that middle school students demonstrated greater competence in the aspects of complexity and perspective-taking, whereas inquiry and skepticism proved to be more challenging. Furthermore, understanding complexity appeared to contribute to better competence in inquiry. In a field-based ecology course focused on a local environmental issue of ethanol production, high school students demonstrated an advancement of their competences on the all aspects of SSR except skepticism (Kinslow, et. al., 2019). In their study on the adaptation of the Quantitative Assessment of Socio-Scientific Reasoning (QuASSR; Romine et al., 2017) into Turkish for high school students, Tüzüngüç and colleagues (2021) reported that students generally exhibited low levels of SSR skills. The findings revealed that while students struggled particularly with the inquiry dimension, they demonstrated their highest competence in the multiple perspective taking dimension. Similarly Öztürk and Roehrig (2025) implemented a STEM focused SSI based instruction on middle school students and found that inquiry and skepticism was challenging aspects for students to develop.

In addition to studies that document learners' competencies in different aspects of SSR separately, there is research highlighting the relationships among SSR dimensions as well as the influence of SSR on learners' decision-making regarding SSI. Romine and colleagues (Romine et al., 2017; Romine, Sadler, Dauer & Kinslow, 2020) have devoted efforts to developing a measurement tool that enable large-scale assessments of learners' competencies across the sub-dimensions of SSR. In these studies, the authors documented that SSR appears to function as a largely unidimensional construct, in that these skills (complexity, multiple perspective, inquiry and skepticism) are closely interrelated and can be regarded as components of a single overarching construct of SSR. In addition, the authors found a hierarchical relationship among the dimensions of SSR, such that recognizing complexity enables students to understand that these issues are difficult to resolve because decision-making is influenced by the vested interests of various parties. Understanding this multiplicity of views on SSI leads students to seek additional information and to adopt a skeptical stance toward the information provided by different stakeholders. This hierarchical relationship suggests that advanced reasoning can be scaffolded by fostering an understanding of multiple perspectives on SSI (Romine et al., 2020). This conclusion is further supported by the study of Ha, Park, and Song (2023), which examined elementary teachers' decision-making regarding COVID-19 school closures and found that preservice teachers' perspectives on the issue strongly influenced the decision-making process. These perspectives shaped how they interpreted and used scientific information, often leading to biased conclusions or decisions.

Overall, the existing literature on the competencies of various groups of learners across the dimensions of SSR has shown that the aspects of complexity and multiple perspectives tend to be relatively less challenging to develop compared to the aspects of

inquiry and skepticism. Additionally, the multiple perspectives dimension appears to play a mediating role in bridging students' understanding of the complexity of SSI with their recognition of the importance of inquiry and skepticism in making informed decisions about such issues. In light of this literature, the present study adopts the conceptual framework of SSR described above to examine preservice science teachers' SSR competencies, based on the premise that without such competencies, it is unrealistic to expect teachers to effectively integrate them into their instruction.

METHOD

Research Context and Participants

Seventy-six third and fourth year preservice science teachers from a public university located in Western Black Sea region of Türkiye participated in the study. The participating preservice teachers were selected based on the voluntarily participation. At the time of the study, the curriculum to which the participating pre-service science teachers were subjected did not include any specific course that addresses socioscientific issues or aims to develop teachers' socioscientific reasoning skills. However, these pre-service teachers had taken or, in the case of the third-year participants, were taking during the time of the study courses that are not directly related to socioscientific issues and reasoning skills but may contribute indirectly to SSR competence, such as genetics and biotechnology, the nature and history of science, and scientific research methods.

Data Collection and Analysis

As a data collection tool, a measurement instrument consisting of two different socioscientific scenarios with open-ended questions was used. These scenarios focused on the use of antibiotics in poultry farming and the production of genetically modified corn for use as biofuel. The scenarios were structured in a way that allows preservice teachers to demonstrate their socioscientific reasoning (SSR) skills. For example, to assess the "multiple perspectives" dimension of SSR, different stakeholder viewpoints on the issue were included within each scenario. Additionally, each scenario included five questions designed to identify participants' SSR dimensions. In the development of these questions, existing studies that aimed to measure SSR sub-dimensions were used as references, and the questions presented in Table 2 were developed accordingly (Sadler et al., 2007; Romine et al., 2020).

After the measurement instrument was developed, it was sent to a total of three academics, two of whom specialize in socioscientific issues in science education. Expert opinions were requested regarding the extent to which the scenarios and questions could assess preservice teachers' SSR skills. Necessary revisions were made based on the feedback received from the experts. Additionally, to ensure that the target group could comprehend the instrument, interviews were conducted with a small group of preservice teachers to assess the readability and clarity of the tool. These preservice teachers did not participate in the actual data collection.

Table 1. Information about Participants and Scenario

	Total Number	Antibiotic in Poultry	Genetically Modified Corn
Third Year Preservice Teachers	38	15	23
Fourth Year Preservice Teachers	38	19	19

The data collection instrument was administered to the students during a class period, and it was stated that their participation was voluntary. Before distributing the instrument, the students were informed about the purpose of the study. Afterwards, the instrument was handed out. It was emphasized that students should answer only the questions they knew and that providing honest responses was important for the outcomes of the research. The analysis of the data was carried out with the qualitative data analysis method of constant comparative analysis (Glaser & Strauss, 1967; Strauss & Corbin, 1998). The authors first went through separately the responses of participants to get familiar with the data (Figure 1).

The data analysis process continued with both authors independently coding all the data and identifying categories through a constant comparative approach. This process concluded with all the data being independently coded by the two authors. In the subsequent phase, the authors met to compare their codes and discussed any discrepancies, resulting in the final version of the coding rubric. This final rubric was then compared with rubrics used in existing studies aiming to assess SSR skills (Sadler et al.,

2007; Romine et al., 2017), and it was determined that it did not fundamentally conflict with the rubrics in the literature. Based on this, the data were reanalyzed using the finalized rubric.

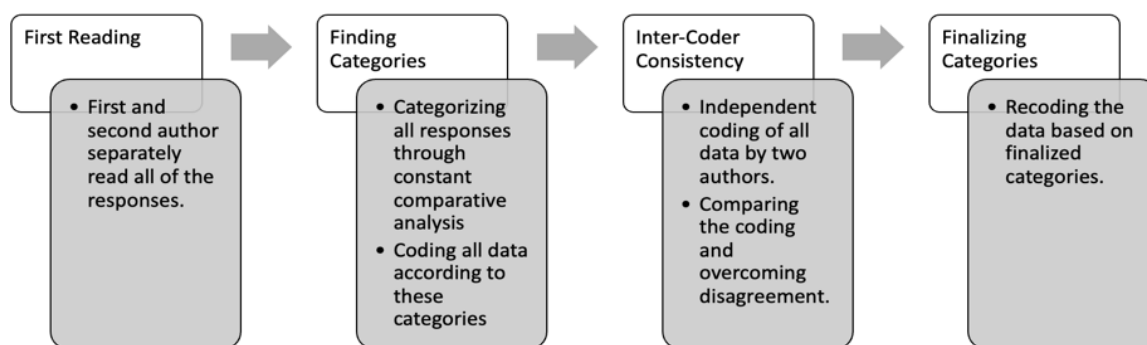


Figure 1. Flow of data analysis process

Table 2. Questions within the scenarios to prompt preservice teachers' SSR

Questions		Purpose of the Questions
Explain the problem related to the use of antibiotics in poultry farming in your own words.	Explain in your own words the problem caused by the use of genetically modified corn in ethanol production.	Making sure that participant understand the issue
Do you think the solution to this problem is difficult? Why?		Assessing the complexity dimension of SSR
Do you think poultry producers who argue that antibiotics are necessary in poultry production and scientific communities who oppose antibiotic use can come up with a common solution to this problem? How?	Do you think producers of non-genetically modified corn and producers of Enogen corn can come up with a common solution to this problem? How?	Assessing the multiple perspective dimension of SSR
If you had to make a decision, do you think the information given in the scenario is sufficient to support or oppose the use of antibiotics in poultry farming? a. If yes: Based on which information in the scenario did you reach your decision and how? b. If no: What kind of information do you think is necessary? What path would you follow to obtain this information?	If you had to make a decision, do you think the information provided in the scenario is sufficient to support or oppose the production of Enogen corn? a.If yes: Based on which information in the scenario did you make your decision and how? b.If no: What kind of information do you think is necessary? What path would you follow to obtain this information?	Assessing the inquiry dimension of SSR
A debate is taking place on a TV program about the use of antibiotics in poultry farming. In this discussion, the head of the poultry producers' union defends the use of antibiotics based on the results of scientific studies. On the other hand, the head of an environmental organization refers to the harmful effects of antibiotics, also based on scientific research. How do you think such opposing views can arise from scientific studies on the same topic?	The ethanol-producing company has stated that it will fund scientists to conduct studies on the problem and share their findings with the public. On the other hand, non-genetically modified corn producers are also working with local scientists to continue research on the issue. Do you think both groups will obtain similar data as a result of their research? Explain.	Assessing the skepticism dimension of SSR

FINDINGS

Our analysis documented that the preservice science teachers generally demonstrated lower levels of competency across the four dimensions of SSR. In particular, they struggled to exhibit higher levels of competency in the inquiry and skepticism dimensions of SSR. When comparing the competencies of the preservice science teachers across different scenarios, it is evident that familiarity with the issue appears to support their ability to reach higher levels of SSR. However, achieving the highest possible level of competency requires dedicated instructional support.

Table 3. Competency levels of complexity dimension of SSR

SSR Dimension	Level 1	Sample Quote
Complexity	The student at this level does not consider the problem difficult to solve. The proposed solution is simple and focuses on a single variable.	"I don't think solving this problem is difficult. Places involved in poultry farming should be kept under control to reduce the use of antibiotics. At the same time, people should be made aware of the harmful effects of antibiotic use. It should also be emphasized that while the feed given to chickens may be beneficial, it can also cause many harms."
	Level 2	Sample Quote
	The student at this level states that the problem is difficult to solve, but attributes the difficulty to one party making the resolution more challenging. Again, the issue is approached with a focus on a single variable.	"Yes, I think it is difficult because those involved in this business and large companies will continuously prevent it, as the losses would outweigh their profits. The best thing to do here is to raise awareness among people and reduce the demand for these chickens."
	Level 3	Sample Quote
	The student at this level is aware of the complexity of the problem and believes that the solution would require addressing both its social and scientific components.	"Yes, solving this problem is definitely difficult. The reason is that the global population is increasing, and the traditional methods used to feed the world's population are both costly and make it very difficult to produce sufficient food."

Table 3 presents the definitions of the three different levels of competence demonstrated by the preservice teachers in the complexity component, along with examples from student responses. In Figure 2, the percentage distribution of the preservice teachers' competencies related to the complexity component of SSR is shown across different scenarios. These results indicate that preservice teachers generally perceive socioscientific issues as simple problems with easy solutions (approximately 68% at Level 1). It appears that they often believe such problems can be solved through a single-variable approach for example, by controlling the actions of chicken producers. Nevertheless, in the scenario about antibiotic use in poultry farming, 26% of students demonstrated Level 3 competency, meaning they believed that the solution would require a multidimensional approach (economic, social, and scientific). However, in the scenario concerning genetically modified corn, the highest level of competency was not observed.

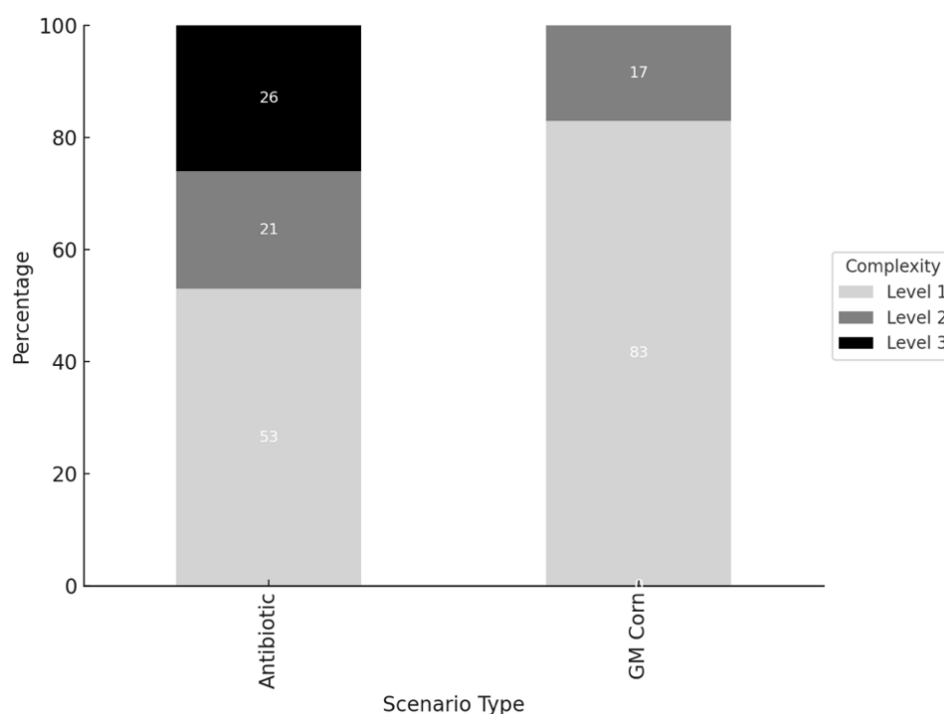
**Figure 2. Percentages of competency levels across scenarios for complexity**

Table 4. Competency levels of perspective taking dimension of SSR

SSR Dimension	Level 1	Sample Quote
Perspective Taking	The student thinks that parties with different opinions can reach an agreement and that this can be achieved in an easy way.	"Yes, for this, chicken producers and the scientific community need to work together. They need to reduce the cost and decrease the use of antibiotics."
	Level 2	Sample Quote
	The student is aware that it is difficult for the different sides of the issue to come to an agreement; however, they attribute this disagreement not to the complexity of the issue itself, but to the resistance of the parties who believe they would be negatively affected by its resolution.	"I don't think chicken producers will find a common solution because, considering these companies and businessmen, I believe they are not even thinking about finding one since their profits and sales would decrease. I think money is more important to them than human health."
	Level 3	Sample Quote
	The student is aware that it is difficult for the different sides of the issue to come to an agreement. Additionally, they point out that the resolution of the issue involves multiple dimensions (such as economic, health-related, etc.).	"I don't think two opposing views will come to a mutual agreement. This is because chicken producers argue that production is not possible without antibiotics, while environmentally conscious individuals claim that antibiotics are harmful. If antibiotics are eliminated, chicken prices will rise; if not, it will lead to biological resistance."

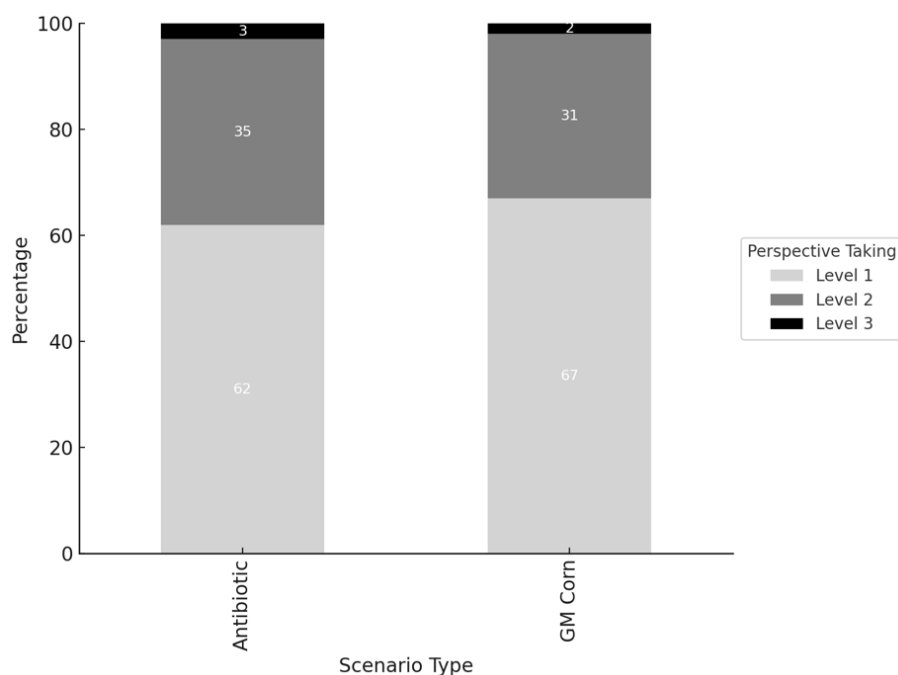
**Figure 3. Percentages of competency levels across scenarios for perspective taking**

Table 4 presents the performance descriptions of the three levels identified for the perspective taking component, along with sample student responses. Figure 3 shows the percentage distribution of preservice teachers' performances related to the perspective taking component across different scenarios. According to these results, the majority of preservice teachers (approximately 65%) demonstrated Level 1 performance in both scenarios. Preservice teachers at this level believe that, despite having differing views, the parties involved in the socioscientific issue can easily come to an agreement and reach a common resolution. Those at Level 2 (approximately 33%) think that the difficulty in resolving the issue stems not from the complexity created by conflicting interests among the parties, but rather from one side's unwillingness to compromise. Unlike the complexity component, no performance differences were observed between scenarios in the perspective taking component.

Table 5 defines the levels of preservice teachers' abilities related to the inquiry component of SSR. Figure 4 presents the percentage of students at these skill levels according to different scenario types. Based on this data, approximately 48% of preservice teachers are observed to be at level 1 in terms of the inquiry component. That is, preservice teachers at this level believe that the information provided in the scenario is sufficient to solve the problem and that no additional information is

needed. A difference is also observed between the scenarios at level 1. In the GM Corn scenario, the percentage of preservice teachers at this level shows a noticeable difference compared to the Antibiotic Use in Poultry scenario. Overall, preservice teachers at level 2 (approximately 47.5%) stated that they needed information focusing on a single dimension of the issue to support better inquiry into solving the problem. At the highest skill level, level 3, only 3% of preservice teachers demonstrated this skill in the Antibiotic Use in Poultry scenario, whereas no preservice teacher reached this level in the other scenario.

Table 5. Competency levels of inquiry dimension of SSR

SSR Dimension	Level 1	Sample Quote
Inquiry	The student believes that there is already sufficient knowledge to solve the issue and that no additional information is needed.	"Yes, it is sufficient. It was explained that a complete ban on the use of antibiotics is not possible. I think it is explained sufficiently."
	Level 2	Sample Quote
	The student believes that there is not enough information to solve the issue. The knowledge they think needs to be questioned focuses only on one component of the problem (such as health or economics).	"No, I don't think there is enough information. I believe the side effects of antibiotics on human health and the problems they may cause should be investigated in more detail. This is about human health."
	Level 3	Sample Quote
	The student believes that there is not enough information to solve the issue. They recognize that the knowledge to be questioned should not be limited to just one component (such as health or economics), but that it is necessary to examine both social and scientific aspects of the issue.	"No. Because both sides need to present objective information to convince each other. WHO should provide data on how antibiotics affect health and to what extent, based on research. On the other hand, chicken producers should present their cost analysis (with and without antibiotics) and their sales percentages."

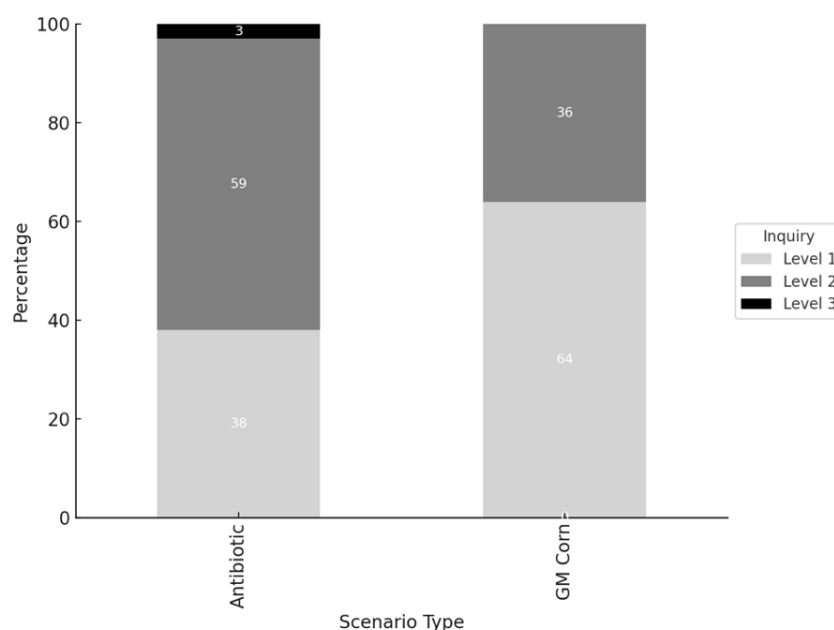
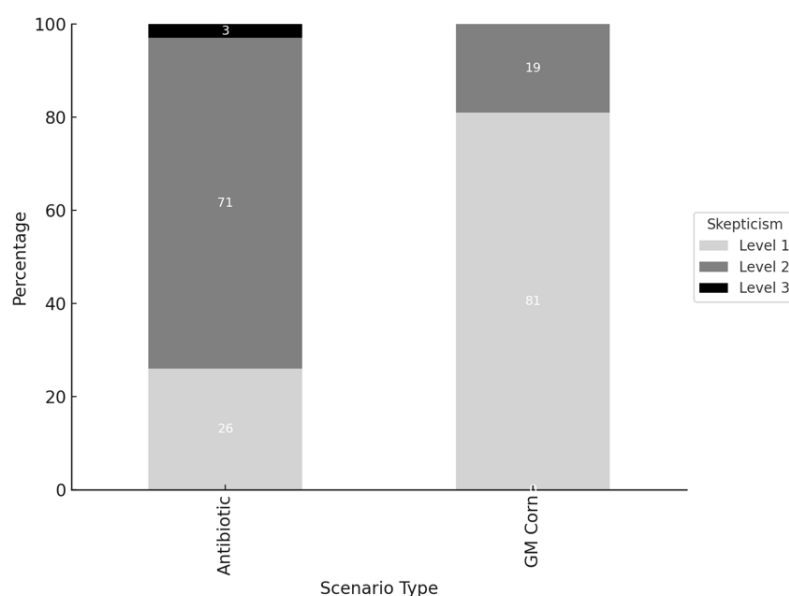


Figure 4. Percentages of competency levels across scenarios for inquiry

Table 6 presents the competency levels and sample responses of preservice teachers regarding the skepticism component of SSR. Figure 5 shows the percentage distribution of preservice teachers' performance levels in the skepticism component based on each scenario. When looking at these results, it is observed that approximately 53.5% of the preservice teachers, regardless of the scenario, demonstrated the level 1 performance. Preservice teachers at this level are not aware of the potential bias that may exist in scientific research. They believe that all scientific studies conducted for the same purpose should yield the same results. Additionally, again independent of the scenario, approximately 45% of the preservice teachers demonstrated level 2 performance. Teachers at this level are aware that researchers may interpret results with bias, yet they still do not consider the possibility that research findings could differ. The highest level of performance for this SSR component was observed in only 3% of preservice teachers, and only in the Antibiotic in Poultry scenario. Another striking finding is the performance difference between scenarios. While most preservice teachers in the Antibiotic in Poultry scenario demonstrated the level 2 competency, most preservice teachers in the GM Corn scenario exhibited the level 1 competency.

Table 6. Competency levels of skepticism dimension of SSR

SSR Dimension	Level 1	Sample Quote
Skepticism	The student believes that as long as scientific studies pursue the same objective, they will not contain opposing views and that the results of scientific research will point to a single reality.	"If the head of the poultry producers and the head of the environmental organization were to research the same issue, I don't think opposing views would emerge. While the poultry producers' head might investigate the advantages of using antibiotics in poultry farming, the environmentalist might research how it affects humans and the environment. In other words, I don't think there would be opposing views when conducting research on the same problem."
	Level 2	Sample Quote
	The student thinks that the differences in scientific studies are related to which aspect of the issue the researchers are associated with.	"Although it may seem necessary to reach the same conclusion from the same data, both sides may highlight and defend the points that benefit them. For example, the president might ignore the more negative aspects because it is economically advantageous for him."
	Level 3	Sample Quote
	At this level, the student questions the reliability of information and considers the potential outcomes that may arise from researcher bias in scientific studies.	"I think the reliability of the sources and scientific research results used by these two opposing sides is important. These opposing views may have emerged through scientific studies. The side supporting poultry farming may conduct research from its own perspective, and the environmentalist side may do the same."

**Figure 5. Percentages of competency levels across scenarios for skepticism**

DISCUSSION, CONCLUSION AND IMPLICATIONS

This study aimed to explore preservice science teachers' SSR competencies through an open-ended, scenario based assessment tool. The results of our analysis show that, in general, preservice science teachers exhibit naive level of SSR competency. They were more competent in appreciating the complexity of SSI compared to other dimensions of SSR. This finding is consistent with prior research studies (Sadler et al., 2007; Öztürk & Roehrig, 2025). Regarding perspective-taking, preservice science teachers were aware of the existence of different perspectives within the issue and were able to identify them. However, they struggled to understand how to balance these challenging aspects to make informed decisions. Instead, they tended to frame this multiplicity of views as a situation in which one side makes the resolution of the issue more difficult.

In terms of the inquiry aspect of SSR, the majority of teacher candidates considered the information provided in the scenarios sufficient to make a decision or requested additional information focusing only on a single aspect of the SSI. This illustrates that the teacher candidates are at an early stage in their process of making informed decisions about SSIs. Similarly, regarding the skepticism aspect, the teacher candidates demonstrated a very limited understanding of how to detect potential bias in the information provided. Prior studies have also highlighted the difficulty of developing these aspects of SSR (Karahana & Roehrig, 2017; Sadler et al., 2011).

Across almost all dimensions of SSR, teacher candidates demonstrated better SSR competency in the context of the Antibiotic Use in Poultry scenario compared to the Genetically Modified Corn scenario. We might attribute this difference to their familiarity

with the issue. Antibiotic use in poultry is a more relatable and familiar topic for preservice teachers compared to genetically modified corn. Ethanol production from genetically modified corn has not been a widely discussed issue in Türkiye, whereas antibiotic use in poultry has been a popular topic, frequently covered on various television channels. Considering this alongside prior studies that demonstrate the impact of deep immersion in the issue on the quality of SSR (Herman, Newton & Zeidler, 2020; Kinslow et al., 2019), we can argue that familiarity with and immersion in the context of an SSI in which learners demonstrate their SSR may have a direct effect on the quality of their reasoning.

The content and delivery of the courses in the current science teacher education program might have a constructive impact on the development of science teacher candidates' reasoning skills regarding socioscientific issues. It is necessary to provide science teacher candidates with learning experiences that can help them develop their reasoning skills regarding socioscientific issues, as well as create opportunities for them to acquire teaching practices through which they can transfer these skills to their future students. We recommend that existing science teacher education curricula in higher education institutions in Türkiye should incorporate courses specifically provide teacher candidates to experience discussions of SSI and improve their SSR competence. This is also important considering that new curricula proposed by Ministry of National Education for K-12, which are currently piloted in some grade levels, are aiming for supporting not only development of conceptual understanding of science concepts but also cultivating science practices as well as student development of virtue, value and action taking toward social issues of local and global issues (MoNE, 2024). SSI based teaching and development of SSR would definitely help cultivating and supporting development of these affective properties in students besides development of conceptual and practical aspects of science.

Moreover, when considering all educational levels in general and specifically within the scope of this study in the context of science teacher education in Turkey, it is observed that there is a significant lack of instructional materials specifically designed for socioscientific issues and proven to be effective, as well as pedagogical approaches that can be implemented during instruction to support the development of socioscientific reasoning skills. Various studies in the literature have shown that SSI-based instructional processes have varying levels of impact on the development of students' SSR skills (Sadler et al., 2011; Kinslow et al., 2019), and have emphasized the importance of investigating the effectiveness of the instructional process. Therefore, there is a need to investigate and develop instructional materials that will support the implementation of socioscientific issues in the classroom, as well as the pedagogical approaches and strategies to be used during the implementation of these materials.

While we employed an open-ended, scenario-based assessment tool for exploring preservice science teachers' socioscientific reasoning (SSR), the applicability of such an assessment to larger participant groups would not be feasible. In addition to that, the limited objectivity in the evaluation process of open-ended questions should also be taken into consideration. Therefore, studies focusing on multiple-choice assessment tools with proven validity and reliability (Romine et al., 2017; Romine et al., 2020; Tüzüngüç et al., 2021) should be increased, and materials to be used in classroom assessment processes should be developed.

Last but not least is to suggest that expansion and/or more detailed specification of the construct of SSR might be needed. Although there are some studies conducted for this purpose (e.g., Newton & Zeidler, 2020), we believe that further investigation of the structure of SSR and clarification of its characteristics would contribute positively both to the instructional process and the development of educational materials.

Declaration of Conflicting Interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

This work was supported by the Scientific Research Projects Coordination Unit of Kastamonu University under grant number KÜ-BAP01/2019-8. Ideas expressed in this material are those of the authors and do not necessarily reflect the views of the funding agency.

Statements of publication ethics

We hereby declare that the study has not unethical issues and that research and publication ethics have been observed carefully.

Researchers' contribution rate

The study was conducted and reported with equal collaboration of the researchers.

Ethics Committee Approval Information

The authors declare that the study was conducted with due considerations of ethical standards. The study was conducted in accordance with the ethical principles approved by the Ethics Committee of Kastamonu University (25.03.2021 decision no:54).

Disclosure Statement

Earlier version of this work was presented at the First International Science, Education, Art and Technology Symposium (2-4 May, 2019) and published its summary in the conference proceedings book.

REFERENCES

- Fowler, S. R., Zeidler, D. L., & Sadler, T. D. (2009). Moral sensitivity in the context of socioscientific issues in high school science students. *International Journal of Science Education* 31(2): 279–296. <https://doi.org/10.1080/09500690701787909>
- Glaser, B. G., & Strauss, A. L. (1967). *The discovery of grounded theory*. Chicago, IL: Aldine.
- Ha, H., Park, W., & Song, J. (2023). Preservice elementary teachers' socioscientific reasoning during a decision-making activity in the context of COVID-19. *Science & Education*, 32, 1869-1886. <https://doi.org/10.1007/s11191-022-00359-7>
- Herman, B. C., Newton, M. H., & Zeidler, D. L. (2020). Impact of place-based socioscientific issues instruction on students' contextualization of socioscientific orientations. *Science Education*, 105, 585-627. <https://doi.org/10.1002/sce.21618>
- Karahan, E., & Roehrig, G. H., (2017). Secondary school students' understanding of science and their socioscientific reasoning. *Research in Science Education* 47 (4): 755–782. <https://doi.org/10.1007/s11165-016-9527-9>
- Kinslow, A. T., Sadler, T. D., & Nguyen, H. T. (2019). Socio-scientific reasoning and environmental literacy in a field-based ecology class. *Environmental Education Research*, 25(3), 388-410. <https://doi.org/10.1080/13504622.2018.1442418>
- Kolstø, S.D., Paulsen, V.H., & Mestad, I. (2024). Critical thinking in the making: students' critical thinking practices in a multifaceted SSI project. *Cultural Studies of Science Education*, 19, 499-530. <https://doi.org/10.1007/s11422-024-10217-3>
- Ministry of National Education [MoNE] (2024). *Science course curriculum for primary and lower secondary schools*. MoNE Publications, Ankara.
- National Research Council (2012). *A framework for K-12 science education: Practices, crosscutting concepts, and core ideas*. Washington, DC: National Academy of the Sciences.
- Newton, M. H., & Zeidler, D. L. (2020). Developing socioscientific perspective taking. *International Journal of Science Education*, 42(8), 1302-1319. <https://doi.org/10.1080/09500693.2020.1756515>
- Öztürk, N., & Roehrig, G. H. (2025). Effects of an integrated STEM unit designed around socioscientific issues on middle schools students' socioscientific reasoning. *International Journal of Science and Mathematics Education*, 23, 1493-1518. <https://doi.org/10.1007/s10763-024-10517-8>
- Patronis, T., Potari, D., & Spiliotopoulou, V. (1999). Students' argumentation in decision making on a socio-scientific issue: Implications for teaching. *International Journal of Science Education*, 21, 745–754. <https://doi.org/10.1080/095006999290408>
- Roberts, D. A. (2007). Scientific literacy/science literacy. In S. K. Abell & N. G. Lederman (Eds.), *Handbook of research on science education*. Mahwah, New Jersey: Lawrence Erlbaum Associates. <https://doi.org/10.1002/tea.21368>
- Roberts, D. A., & Bybee, R. W. (2014). Scientific literacy, science literacy, and science education. In S. K. Abell & N. G. Lederman (Eds.), *Handbook of research on science education* (pp. 545–558). New York, NY: Routledge.
- Romine, W. L., Sadler, T. D., & Kinslow, A. T. (2017). Assessment of scientific literacy: Development and validation of the quantitative assessment of socio-scientific reasoning (QuASSR). *Journal of Research in Science Teaching*, 54 (2): 274–295. <https://doi.org/10.1002/tea.21368>
- Romine, W. L., Sadler, T. D., Dauer, J. M., & Kinslow A. T. (2020). Measurement of socio-scientific reasoning (SSR) and exploration of SSR as a progression of competencies. *International Journal of Science Education*, 42(18), 2981-3002. <https://doi.org/10.1080/09500693.2020.1849853>
- Sadler, T. D. (2004). Informal reasoning regarding socioscientific issues: A critical review of research. *Journal of Research in Science Teaching*, 41, 513–536. <https://doi.org/10.1002/tea.20009>
- Sadler, T., & Zeidler, D. (2005). Patterns of informal reasoning in the context of socioscientific decision making. *Journal of Research in Science Teaching*, 42 (1), 112–138. <https://doi.org/10.1002/tea.20042>
- Sadler, T. D., Barab, S., & Scott, B. (2007). What do students gain by engaging in socioscientific inquiry? *Research in Science Education*, 37(4), 371–391. <https://doi.org/10.1007/s11165-006-9030-9>
- Sadler, T., D., & Zeidler, D., L. (2009). Scientific literacy, PISA, and socioscientific discourse: Assessment for progressive aims of science education. *Journal of Research in Science Teaching* 46 (8): 909–921. <https://doi.org/10.1002/tea.20327>
- Sadler, T. D., Klosterman, M. L., & Topcu, M. S. (2011). Learning science content and socio-scientific reasoning through classroom explorations of global climate change. In T. D. Sadler (Ed.), *Socio-scientific issues in the classroom: Teaching, learning and research* (pp. 45–77). Springer.
- Strauss, A., & Corbin, J. (1998). *Basics of qualitative research: Techniques and procedures for developing grounded theory*. Thousand Oaks, CA: Sage.
- Tüzüngüç, B., Doğan, Ö., & Han Tosunoğlu, Ç. (2021). Adaptation of socio-scientific reasoning scale to Turkish: A validity and reliability study. *Electronic Journal of Social Sciences*, 20 (79), 106-1078. <https://doi.org/10.17755/esosder.763534>

- Zeidler, D. L., & Sadler, T. D. (2010). An inclusive view of scientific literacy: Core issues and future directions of socioscientific reasoning. In Linder, C., Östman, L., Roberts, D. A., Wickman, P., Ericksen, G. & MacKinnon, A. (Eds). *Exploring the Landscape of Scientific Literacy*. (pp. 176-192). Taylor & Francis: New York.
- Zeidler, D., L., Herman, B. C., Ruzek, M., Linder, A., & Lin., S. (2013). Cross-cultural epistemological orientations to socioscientific issues. *Journal of Research in Science Teaching* 50 (3): 251–283. <https://doi.org/10.1002/tea.21077>
- Zeidler, D.L., & Sadler, T.D. (2023). Exploring and expanding the frontiers of socioscientific issues: Crossroads and future directions. In N. G. Lederman, D.L. Zeidler, & J.S. Lederman (Eds.), *Handbook of Research on Science Education, Volume III* (pp. 899-929). Routledge.
- Zohar, A., & Nemet, F. (2002). Fostering students' knowledge and argumentation skills through dilemmas in human genetics. *Journal of Research in Science Teaching* 39 (1): 35–62. <https://doi.org/10.1002/tea.10008>

Annexes

Scenario 1: Antibiotics in Poultry

A Frightening Danger: Increasing Antibiotic Resistance

The World Health Organization (WHO) identifies increasing antibiotic resistance as one of the most significant threats to global health and food safety. Antibiotic resistance is defined as the ineffectiveness of antibiotics due to bacteria developing defenses as a result of the excessive and incorrect use of antibiotics. It is estimated that approximately 70–80% of the total amount of antibiotics used globally is utilized in livestock farming. In light of this information, the use of antibiotics in livestock plays a major role in the development of antibiotic resistance. Furthermore, scientific studies have shown that antibiotic resistance developing in livestock negatively affects humans as well.

Poultry farming holds a significant place in meeting people's protein needs both in our country and around the world. In industrial poultry farming, antibiotics are used for various purposes. These include preventing infections, treating diseases, and promoting the growth of chickens. Since 2006, the use of antibiotics to promote the growth of livestock has been banned in our country. However, antibiotics are still used under veterinary supervision for the treatment and prevention of disease by mixing them into the feed or water of chickens.

There are ongoing serious debates about the use of antibiotics in poultry farming. Some scientific circles argue that the use of antibiotics in poultry should be completely banned, regardless of the reason, because antibiotic use causes bacteria to develop resistance, and there is a risk that these resistant bacteria may be transmitted to humans directly or indirectly. In such a case, antibiotics would lose their therapeutic effect, and incurable diseases caused by resistant bacteria may emerge in people who consume these chickens.

Poultry farmers, on the other hand, argue that a complete ban on antibiotic use is not feasible and that raising chickens without antibiotics would not be an economically viable option. In other words, they state that raising chickens without antibiotics would be more costly and would therefore lead to an increase in chicken prices. As a result, they claim that people's access to chicken meat would be limited.

In street interviews conducted on this topic, people express that they do not want to consume chickens raised with antibiotics. However, when reminded of the farmers' view that raising antibiotic-free chickens is not an economically viable solution, they state that chicken prices are already high and that any further increase would put the public in a difficult position.

Scenario 2: Genetically Modified Corn

Alternative Energy Source: Ethanol

Today, interest in alternative sources is steadily increasing in order to meet the growing energy demand. One of these alternative energy sources is ethanol, which can be produced from various grains. The United States and Brazil, which account for 70% of global ethanol production, obtain the majority of their ethanol by processing corn.

The enzyme alpha-amylase plays an important role in increasing the efficiency of ethanol production. Since 2017, farmers in North Dakota have been growing a genetically modified corn called Enogen for the ethanol industry. Enogen corn contains high levels of this enzyme.

A representative from one of the ethanol-producing companies states that, in the past, alpha-amylase had to be purchased and added to the mixture during ethanol production. Thanks to Enogen corn, there is no longer a need to purchase this enzyme externally, as the corn itself contains it in abundance. Additionally, farmers who produce Enogen corn are paid an extra fee per bushel (a unit used to measure grain volume), which benefits the farmers financially.

Farmers who produce non-genetically modified corn oppose the production of Enogen corn. Their primary concern is the risk of genetic contamination from fields growing Enogen corn to their own fields due to environmental factors. Tests conducted by farmers producing non-GMO corn have found traces of Enogen corn in the flour produced from their own harvests. These findings confirm the farmers' concerns. Due to this contamination, the high concentration of alpha-amylase in Enogen corn causes the starch in the corn to break down. As a result, corn chips become crumbly and corn flakes lose their texture. According to the North American Millers' Association, this issue can occur even if only 1 Enogen corn kernel is mixed with 10,000 kernels of food-grade

corn. While the ethanol industry benefits from Enogen corn, it is argued, especially by farmers and millers, that this type of corn severely undermines the food industry.

Table. Competency levels of sub-dimensions of SSR (Genetically Modified Corn Scenario)

SSR Dimension	Level 1	Sample Quote
Complexity	The student at this level does not consider the problem difficult to solve. The proposed solution is simple and focuses on a single variable.	"I don't think solving this problem is difficult because these genetically modified corns can be grown collectively and in a protected area without contamination in different locations."
	Level 2	Sample Quote
	The student at this level states that the problem is difficult to solve, but attributes the difficulty to one party making the resolution more challenging. Again, the issue is approached with a focus on a single variable.	"I think it is difficult because Enogen corn kernels have been found through contamination even in the corn of farmers who do not use Enogen corn. All genetically modified corn must be destroyed. But that is very difficult because the seeds have spread to many places. Even if I don't use it, there will be contamination of my products by someone nearby who does. This will be a challenging process."
Perspective Taking	Level 3	Sample Quote
	The student at this level is aware of the complexity of the problem and believes that the solution would require addressing both its social and scientific components.	This level of competence were not seen in the responses of the participants.
	Level 1	Sample Quote
Inquiry	The student thinks that parties with different opinions can reach an agreement and that this can be achieved in an easy way.	"Yes, they can produce it. It can be grown in a different and protected area. This way, when production is done in different locations, the possibility of contamination due to environmental effects would be eliminated. This would be a solution."
	Level 2	Sample Quote
	The student is aware that it is difficult for the different sides of the issue to come to an agreement; however, they attribute this disagreement not to the complexity of the issue itself, but to the resistance of the parties who believe they would be negatively affected by its resolution.	"I think they cannot produce it because while genetically modified Enogen corn producers have benefited, the non-genetically modified corn producers have suffered losses."
Inquiry	Level 3	Sample Quote
	The student is aware that it is difficult for the different sides of the issue to come to an agreement. Additionally, they point out that the resolution of the issue involves multiple dimensions (such as economic, health-related, etc.).	"I don't think they can come up with a common solution because the ethanol industry that benefits from Enogen corn will continue to use it in order not to import ethanol. Regular corn producers will oppose Enogen corn because it will cause deterioration in the structure of natural corn and reduce its quality. Therefore, I don't think they can come up with a common solution."
	Level 1	Sample Quote
Inquiry	The student believes that there is already sufficient knowledge to solve the issue and that no additional information is needed.	"I think it can be considered sufficient because statements such as the benefits and harms of Enogen corn production, where it is mostly produced, and which enzyme increases its efficiency are included."
	Level 2	Sample Quote
	The student believes that there is not enough information to solve the issue. The knowledge they think needs to be questioned focuses only on one component of the problem (such as health or economics).	"I don't think so. Because only superficial information was given. I would have liked more information about Enogen corn. I would have liked the benefits and harms to be presented more clearly. A decrease in product quality was observed in those who used Enogen corn. I would have liked research on factors that could improve this and clear information to be provided."
Inquiry	Level 3	Sample Quote
	The student believes that there is not enough information to solve the issue. They recognize that the knowledge to be questioned should not be limited to just one component (such as health or economics), but	This level of competence were not seen in the responses of the participants.

SSR Dimension	Level 1	Sample Quote
	that it is necessary to examine both social and scientific aspects of the issue.	
	Level 1	Sample Quote
	The student believes that as long as scientific studies pursue the same objective, they will not contain opposing views and that the results of scientific research will point to a single reality.	"I think they will obtain similar data because the producers using Enogen corn will question the poor quality of the products. Other producers will question the effects that damage their products, and similar results will emerge."
	Level 2	Sample Quote
Skepticism	The student thinks that the differences in scientific studies are related to which aspect of the issue the researchers are associated with.	"I don't think they will reach a common decision because there are scientists supporting both views. Both sides will defend their own opinions. I think the research results will differ. Everyone will conduct research that serves their own interests."
	Level 3	Sample Quote
	At this level, the student questions the reliability of information and considers the potential outcomes that may arise from researcher bias in scientific studies.	This level of competence were not seen in the responses of the participants.



Understanding What You Read From İhsâu'İ Ulûm to Modern Education: Turkey's PISA Test Performance

İhsâu'İ Ulûm'den Modern Eğitime Okuduğunu Anlamak: Türkiye'nin PISA Testi Performansı

Gözde ÇAVDAR¹

Keywords

1. Reading Comprehension
2. Fârâbî
3. Linguistics
4. PISA tests
5. Modern education

Anahtar Kelimeler

1. Okuduğunu anlama
2. Fârâbî
3. Dilbilim
4. PISA testleri
5. Modern eğitim

Received/Başvuru Tarihi

21.05.2025

Accepted / Kabul Tarihi

10.10.2025

Abstract

Purpose: The aim of this study is to reveal the deficiencies in students' reading comprehension skills in the Turkish education system and to evaluate the importance of linguistics in Farabi's İhsâu'İ Ulûm from the perspective of modern education.

Design/Methodology/Approach: In the research, Farabi's "İhsâu'İ Ulûm", PISA test results and various academic studies were used. By applying literature review and text analysis method, the data were evaluated with theoretical comparison and interpretation techniques.

Findings: It was found that students in Turkey scored below the OECD average in PISA tests in terms of reading comprehension ability. The main reasons for this situation were found as school type, socio-economic status, inadequacy of educational materials and discontinuity in the education system. It was also observed that al-Farabi considered the science of language as the basis of all sciences and emphasized that expressing and understanding thoughts correctly requires mastery of language rules. The views of al-Farabi, stating that grammar is prioritized over logic in education, serve as a guide in solving modern educational problems.

Highlights: In order to improve students' reading and comprehension skills, it is necessary to increase access to materials, update educational content with a focus on analysis and interpretation, and ensure continuity and stability in educational policies.

Öz

Bu çalışmanın amacı, Türk eğitim sisteminde öğrencilerin okuduğunu anlama becerilerindeki eksiklikleri ortaya koymak ve modern eğitim perspektifinden Farabi'nin İhsâu'İ Ulûm'unda dilbilimin önemini değerlendirmektir.

Materyal ve Yöntem: Araştırmada Farabi'nin İhsâu'İ Ulûm adlı eseri, PISA test sonuçları ve çeşitli akademik çalışmalar kullanılmıştır. Literatür taraması ve metin analiz yöntemleri uygulanarak, veriler yorum tekniğiyle değerlendirilmiştir.

Bulgular: Türkiye'de öğrencilerin PISA testlerinde okuduğunu anlama becerisi açısından OECD ortalamasının altında puan aldığı görülmüştür. Bu durumun başlıca nedenleri okul türü, sosyo-ekonomik durum, eğitim materyallerinin yetersizliği ve eğitim sistemindeki süreksizlik olarak belirlenmiştir. Ayrıca, Farabi'nin dil bilimini tüm bilimlerin temeli olarak gördüğü ve düşüncelerin doğru ifade edilip anlaşılmasının dil kurallarına hâkimiyeti gerektirdiğini vurguladığı gözlemlenmiştir. Farabi'nin, eğitimde mantıktan önce grameri önceliklendiren görüşleri modern eğitim sorunlarının çözümünde yol gösterici niteliktedir.

Önemli Vurgular: Öğrencilerin okuma ve anlama becerilerini geliştirmek için materyallere erişimin artırılması, eğitim içeriklerinin analiz ve yorum odaklı olacak şekilde güncellenmesi ve eğitim politikalarında süreklilik ve istikrarın sağlanması gerekmektedir.

¹ Corresponded Author, Kastamonu University, School of Foreign Languages, Kastamonu, Türkiye; <https://orcid.org/0000-0003-4748-1342>

INTRODUCTION

Knowledge has remained the most significant source of power for all communities throughout history. From the most primitive societies to today's advanced civilizations, people have utilized knowledge both to survive and to secure a place in the competitive atmosphere of international platforms, making tremendous efforts to expand existing knowledge. When evaluated in terms of its nature, knowledge today is categorized into natural knowledge, scientific knowledge, philosophical knowledge, artistic knowledge, and religious knowledge (Mengüşoğlu, 2015: 74–93). However, due to its role as the fundamental means of relating to external objects, it has occupied a central place in the agenda of many philosophers from Ancient Greece to the present—most notably Plato and Aristotle—who categorized sciences in hierarchical order to reach the good and the beautiful as an outcome of the state of knowing. The aim of this study is to establish a relationship between al-Fârâbî's emphasis on linguistics and the reading comprehension skills of Turkish students who participated in the PISA tests.

Although neither Plato nor Aristotle explicitly discussed language, al-Fârâbî, who brought Islamic culture and Ancient Greek philosophy together, assigned significant importance to the science of language. In his work *Ihsâu'l Ulûm* (The Classification of the Sciences), where he classified sciences, al-Fârâbî prioritized linguistics over even logic, assigning it to the first rank, while placing logic afterward. From this perspective, the study focuses on the ability to comprehend written texts in a particular language, as this competence plays a central role in both contributing to a culture's internal richness and ensuring alignment with global scientific, cultural, and economic developments.

The PISA test is an international assessment tool conducted every three years since 2000 by the Organisation for Economic Co-operation and Development (OECD). It evaluates 15-year-old students' proficiency in reading, mathematics, and science—skills deemed essential for full participation in social and economic life. The most fundamental way for Turkey to play an active role in the competitive environment that includes other countries is to evaluate its education system and implement necessary improvements, thereby raising its children to the level of those from globally dominant countries—or even beyond. Therefore, analyzing PISA test results is of considerable importance. This study focuses particularly on the 2009 and 2018 PISA tests, as reading literacy was designated the primary focus in those years. Since Turkey joined PISA in 2003, there is no dataset available for the 2000 cycle. Regrettably, Turkish students consistently lag their international peers, especially in the reading comprehension category (OECD, 2023: 1). This indicates a deficiency in the Turkish education system's capacity to adequately train students in reading comprehension. It is evident that current modern education strategies in Turkey fall short in terms of student success. However, no definitive conclusions have been reached regarding the methodological parameters causing this shortfall. In the subsequent sections, this issue will be discussed considering relevant literature and PISA report data.

Classification of Sciences

The classification of sciences began with Plato in Ancient Greek philosophy and continued with Aristotle's divergent interpretation. In his work *The Republic*, Plato discusses the theory of knowledge, stating: "Those who study scientific subjects must not rely on their senses, but must use their reasoning. However, since they examine these subjects based on assumptions without reaching first principles, they cannot truly understand objects that can only be comprehended through such principles" (Plato, 2010: 551d). Through this statement, Plato emphasizes the importance of reaching knowledge, while also providing clues about how knowledge can be attained and how it should be defined. In the seventh book of *The Republic*, Plato (2010) concentrates on the sciences that should be taught to those who will rule the state, presenting what is considered the first example of scientific classification in history (Erdoğan, 2009: 138). In this dialogue, Plato regards as sciences only those disciplines that are entirely intellectual in nature, devoid of empiricism, and yield conclusions through deductive reasoning. These sciences, in order, are arithmetic, geometry, astronomy, harmony, and above all, dialectics, which encompasses and transcends all others. Plato asserts that the subject matter of true knowledge or science (*episteme*) is the realm of true being and the intelligible world. He claims that if the object of investigation is a visible or perceptible entity, then true knowledge cannot be attained (Plato, 2010: 529b). In this regard, science concerns what is immutable, non-particular, and non-sensory—things that cannot be otherwise. Science is, therefore, that which is rational and unchanging. Conversely, any domain associated with the world of appearances is inherently subject to change, transformation, coming-to-be, and decay. As a result, human beings develop opinions (*doxa*) about such matters. The objects of thought, however, are products of knowledge and belong to the domain of being. On this point, Plato states that no object can simultaneously belong to both *episteme* and *doxa*, emphasizing that *doxa* refers to a state between knowledge and ignorance (Plato, 2010: 478d).

Thus, anything that does not reach a first principle, which is individual, subject to change, and capable of being otherwise cannot be regarded as true science. Additionally, based on his famous Allegory of the Cave, Plato considers the sensible world entirely as a world of shadows. His ultimate objective is to attain knowledge, and he views the world merely as a path toward achieving that knowledge. For this reason, he completely excludes empirical investigations of the physical world (Afşar, 1997: 149). Consequently, within Plato's classification of sciences, there is no place for the so-called natural sciences, which are grounded in human experience (Arslan, 2006: 237). According to Plato, disciplines such as physics, chemistry, biology, and medicine are not recognized as true sciences.

When approached from the perspective of Plato's student, Aristotle, it becomes evident that Aristotle did not follow Plato's path but rather developed a significantly different perspective on the classification of sciences. Aristotle argued that the desire to know is innate and that its indication lies in the pleasure we derive from sensory experiences—even at the most basic level and beyond mere utility (Aristotle, 2009: 980b). In this regard, Aristotle's position stands in direct contrast to that of his teacher, Plato, particularly in relation to the role of the senses and opinions. For instance, while Plato emphasized that true knowledge consists of universal knowledge situated at the level of first principles, Aristotle maintained that those who possess only theoretical universal knowledge lack experience. Because they are not acquainted with the underlying causes of universals, they are unable to reach specific conclusions. Thus, Aristotle argues that particulars have an independent existence (Aristotle, 2009: 981b). From this perspective, it would be appropriate to position Aristotle as a pioneer of the realist approach, which emphasizes causal explanation through the discovery of particulars (Keat & Urry, 2001: 18). Given his emphasis on experiment and observation, it would not be incorrect to regard Aristotle as a philosopher who made significant contributions to the empiricist and practical approach. In contrast to Plato's approach, which accepts only sciences that deal with ideals or universals within a theoretical framework, Aristotle evaluated sciences as diverse fields of being, each with distinct goals, starting points, and domains. This perspective (Cevizci, 2016: 393) brought about major differences in the classification of sciences.

From this point of view, Aristotle developed his own unique categorization of the sciences, evaluating each science within its own framework, and further identifying the relationships and distinctions between philosophy and the sciences (Erdoğan, 2009: 153). According to him, "each science investigates certain principles and causes related to its own subject matter" (Aristotle, 2009: 1064a). Based on this understanding, it is not difficult to comprehend why Aristotle included in his classification disciplines such as medicine, physics, chemistry, and biology—fields that can be characterized as practical sciences. Aristotle's classification of the sciences is fundamentally structured into three categories: Theoretical (*Theoretikē*), Practical (*Praktikē*), and Productive (*Poiētikē*) sciences (Reeve, 2009: 15). "The primary aim of each category is to seek knowledge, but their end goals vary; acquiring knowledge, taking action, and creating useful and beautiful objects" (Ross, 2005: 21). Although this aim of reaching the 'good' resembles the final cause found in Plato's thought, a closer look at the specific disciplines Aristotle places under each category reveals a clear divergence in intellectual orientation.

Accordingly, under Theoretical Sciences, Aristotle includes disciplines such as theology, philosophy, mathematics, and the natural sciences, which pursue knowledge for its own sake, without any practical end. Practical Sciences, which treat knowledge as a guide to action, include ethics, household management, political governance, and politics. Under the Productive Sciences—those aimed at producing something beneficial or beautiful—he places the applied sciences and crafts, such as medicine and architecture (Reeve, 2009: 15). From this perspective, while Aristotle's valuation of mathematics differs in its priority compared to Plato's, the discipline itself retains its significance for both philosophers. However, the same cannot be said for the natural sciences such as physics, biology, and chemistry. While Plato excluded these from the domain of true sciences, Aristotle emphasized their importance and included them in his classification. Additionally, Aristotle did not regard logic as an independent science. Instead, he treated it as a general intellectual discipline that must be employed by anyone investigating any science (Ross, 2005: 21). Indeed, at the core of all sciences lies the activity of reasoning about particulars, making judgments, and interpreting them through the lens of logic.

During the Islamic Golden Age, al-Fârâbî (870–950), a prominent Turkish-Islamic philosopher known by the honorific title "The Second Master" (*al-Mu'allim al-Thânî*), undertook the study and translation of the works of influential Ancient philosophers such as Plato and Aristotle, and sought to establish a connection between Islam and philosophy by imbuing Islamic thought with a philosophical dimension (Farabî, 2019: xvii). Additionally, al-Fârâbî is considered the founder of political philosophy in the medieval Islamic world and holds a unique position for his assertion that "in order for a nation to reach the pinnacle of its cultural history, it must examine the stages of development in language and thought" (Korkut, 2015: 26). For these reasons, al-Fârâbî is a philosopher worthy of careful examination, not only for the depth of his philosophical work but also for his pivotal role in introducing Ancient Greek philosophy to the Islamic world during a period described as the Golden Age—an era shaped by an intense emphasis on knowledge and science.

Like Plato and Aristotle, whose works he translated and through which he served as a bridge between Ancient Greek philosophy and the Islamic world, al-Fârâbî also classified the sciences. In his book *Ihsâu'l Ulûm (The Classification of the Sciences)*, which contains this classification, al-Fârâbî states that his purpose in writing the work is to list the sciences of his time one by one and to explain into which categories each science is divided internally (Farabî, 2019: 1). Furthermore, while introducing the book, he emphasizes that those who consult it will understand what discipline they are engaging with, what benefits it will provide to them, and which virtue they will ultimately attain as a result (Farabî, 2019: 1). Al-Fârâbî, like Plato and Aristotle, viewed studying the sciences as essential to attaining higher virtue and clearly expressed this belief.

In his work *Ihsâu'l Ulûm (The Classification of the Sciences)*, al-Fârâbî classified the sciences into five main groups: the Science of Language, the Science of Logic, the Mathematical Sciences, the Natural and Theological Sciences (Physics and Metaphysics), and the Sciences of Politics, Jurisprudence, and Theology. Notably, among these disciplines, the Science of Language—absent from Aristotle's classification—occupies the foremost position in al-Fârâbî's ordering. Moreover, logic, which Aristotle regarded merely as a methodological tool and not as an independent science, is placed by al-Fârâbî immediately after the Science of Language, in the second position (Farabî, 2019: xi). This reveals that, for al-Fârâbî, the Science of Language is the most fundamental and indispensable of all sciences, and its most essential companion in achieving success is the Science of Logic.

al-Fârâbî and the Science of Language in *Ihsâu'l- Ulûm*

When enumerating the sciences, al-Fârâbî places the Science of Language at the very top of the list—an order that is by no means coincidental. According to al-Fârâbî (2019: 3), language underlies every science. The fundamental content of the Science of Language, he states, is “the learning of utterances that signify meaning [...] and knowledge of the rules governing these utterances.” In this context, it becomes possible to derive many expressions by memorizing a limited set of linguistic units (Farabî, 2019: 4). This, in turn, implies the application of logic. Although the language whose rules are articulated in the text is Arabic, al-Fârâbî’s account of language pertains to universal linguistic principles and is adaptable to the languages of all nations. As al-Fârâbî himself states (2019: 15), just as there are common linguistic rules across languages, each community also possesses its own unique linguistic characteristics. In his view, knowing and internalizing the language of a society or nation is a prerequisite for being considered a member of that society or nation (Korkut, 2015: 27). Indeed, language is the precondition of communication, and individuals can only contribute to a shared culture if they converge upon a common linguistic medium.

al-Fârâbî (2019: 4) divides the Science of Language into seven parts: the Science of Isolated Words, the Science of Compound Words, the Science of the Rules Governing Words in Isolation, the Science of the Rules Governing Words in Combination, the Rules of Correct Writing, the Rules of Correct Reading, and the Science of Poetic Rules. Al-Fârâbî provides general principles for each of these linguistic divisions, and his treatment of the *Rules of Correct Reading* is especially notable for its detailed consideration—even extending to punctuation marks. Although each of these sections may appear to focus solely on the structural elements of language, al-Fârâbî (2019: 10–16) establishes a direct connection between grammar and logic. He argues, particularly in the section on logic, that the correct expression of thought—and indeed the accurate understanding of expressed thought—requires prior knowledge of linguistic rules, and especially grammar. For al-Fârâbî, grammar is not merely a technical subject but a necessary foundation for rational discourse and comprehension.

In *Ihsâ' al-'Ulûm*, al-Fârâbî (2019: 10) discusses the similarities between the Science of Language and the Science of Logic, noting that just as we reach knowledge related to reason and rational entities through the art of logic, we likewise access the general rules of speech through the art of language. It is also possible to observe that al-Fârâbî establishes connections between grammar and logic in his other works beyond *Ihsâu'l Ulûm* (as cited in Korkut, 2015: 27). For instance, in his treatise *Advice on the Path to Happiness*, al-Fârâbî emphasizes that the discipline which enables a person to correctly express what they think is not logic, but grammar (as cited in Korkut, 2015: 27). In this regard, al-Fârâbî’s meticulous approach to language—and, by extension, to expression and the interpretation of language—renders his perspective particularly significant for the present study.

METHOD/MATERIALS

Research Design

This study is designed as a qualitative research project. In qualitative research, the inclusion of the researcher’s subjective interpretation emerges as a necessity for the research process (Denzin et al., 2024: 3). The study employs the document analysis method, and the data were evaluated through content analysis techniques.

Data Collection Tools

The main axis of the study consists of analyzing and correlating *Ihsâu'l Ulûm (The Classification of the Sciences)*—a seminal work by al-Fârâbî, one of the most prominent philosophers of the Islamic-Turkish intellectual tradition—and the PISA test reports published by the OECD. Accordingly, *Ihsâu'l Ulûm* as well as the PISA 2009, PISA 2018, and PISA 2022 educational reports were selected for analysis. Since the PISA 2009 and PISA 2018 tests designated reading literacy as the core domain, the study focuses particularly on those two cycles in greater detail.

FINDINGS

PISA Tests and Turkey’s Reading Comprehension Performance

Turkey’s education system is characterized by the continuous evaluation of students’ knowledge. From the moment they begin formal education, students are subjected to numerous examinations in various subjects and internalize the educational system’s emphasis on the practice of “recalling acquired knowledge” to secure a good job—and by extension, a good life. Considering the general structure of the system, these exams typically take the form of reading written texts and answering a specific number of questions within a limited time frame. This structure underscores the importance of assessing the extent to which students comprehend what they read and how effectively they can recall that information. These evaluations have now extended beyond the national framework, being conducted through international projects established within the context of an increasingly globalized world order. In this regard, Turkey participates in initiatives such as the Progress in International Reading Literacy Study (PIRLS), the Trends in International Mathematics and Science Study (TIMSS), and the Programme for International Student Assessment (PISA), all of which serve to identify and address existing shortcomings in the national education system.

The PISA test is an international assessment administered every three years by the Organisation for Economic Co-operation and Development (OECD), of which Turkey has been a founding member since 1961. It aims to measure the extent to which 15-year-old students can apply the knowledge they have acquired at school to real-life situations. The test primarily evaluates

students' competencies in reading, mathematics, and science. Since its inception in 2000, each cycle of the test has focused in depth on one of these three core skill areas. According to information available on the official OECD website, Turkey began participating in the PISA test in 2003, and among the three core domains, reading literacy was selected as the primary focus in the years 2000, 2009, and 2018 (MEB, 2022: 3). Official data indicate that Turkey's average reading score in the 2009 PISA test was 464 (OECD, 2010: 15). In this test, it was reported that the average performance level of Turkish students was at Level 2. At this level, students are expected to:

“Locate information that meets multiple conditions, make comparisons or contrasts based on a single clearly defined feature, interpret the meaning of a well-defined part of a text—even if the information is not explicitly stated—and make connections between the text and their own experiences” (OECD, 2010: 52).

This level is regarded in the PISA report as the basic reading proficiency necessary for students to participate effectively and productively in everyday life. Level 6 on the PISA scale shows that students can overcome their biases when they encounter new, unexpected information. At this level, students are expected to: “Distinguish both explicitly stated and more subtly presented information in a text; and evaluate that information critically, drawing on sophisticated understandings that extend beyond the text itself” (OECD, 2010: 49). Among Turkish students who participated in the PISA 2009 test, almost none were reported to have achieved performance at Level 6. This indicates that students in Turkey were unable to move beyond the basic level of reading proficiency in the PISA reading category in 2009.

In the 2018 PISA test, the average reading performance of students in Turkey was reported as 466 (OECD, 2019: 17). Among the 79 participating countries, the average reading score was 453, placing Turkey above the global average and ranking 40th overall. However, the average reading score among the 37 OECD member countries was 487, and Turkey ranked 31st among them. In the same test, the distribution of Turkish students across proficiency levels was as follows: 26.1% performed below Level 2, 30.2% at Level 2, 26.9% at Level 3, 13.5% at Level 4, 3.1% at Level 5, and only 0.2% at Level 6 (OECD, 2019: 210). While students demonstrated improvement compared to previous tests, they still failed to exceed the basic proficiency threshold in reading literacy. The fact that Turkish students mostly performed at Level 2 is particularly notable in light of Ahmet Benzer's (2019) study, which examined the relationship between Turkish language textbooks and PISA tests. According to his findings, most reading texts in Turkish textbooks were aligned with the cognitive processes of Levels 1b (29%), 2 (25%), and 1a (24%) (Benzer, 2019: 101). In this context, it is not surprising that Turkish students performed above the OECD average specifically at Level 2 in the 2018 PISA reading literacy assessment (OECD, 2019).

DISCUSSION

In both years evaluated, it is evident that the average performance of Turkish students in the PISA reading literacy assessment remained below the overall OECD average. For example, PISA 2009 test (OECD, 2010) reveals that Turkish students' reading comprehension level fell behind the OECD average with 464 points. In this test, none of the students succeeded in level 6 and the average of level 5 succeeders were 1,8 which was one-fourth of OECD countries' accomplishment average at this level. The data obtained from the test reports suggest that students in Turkey are insufficiently equipped in reading comprehension, and despite showing some progress over time, their performance continues to linger at a basic level. This, in turn, may negatively affect their achievement in other academic disciplines. Notably, the logic categorized in al-Fârâbî classification of the sciences is, after all, the language of universal thought and reasoning. It should be acknowledged that students can enhance their capacity to distinguish between meaningful ideas and adapt what they read to their social lives only with the support of their families and the educational system.

The scores obtained from these tests can be correlated with various components of Turkey's education system. For instance, in her study on the factors influencing reading achievement in the PISA 2018 test, Özlem Okatan (2021: 332–333) considered variables such as school type, student-teacher ratio, material shortages, school location, socio-economic status, gender, and pre-school education. According to her findings, all variables except the student-teacher ratio were significantly related to reading performance (Okatan, 2021: 337–346). Students attending Science High Schools had an average reading score of 581.37, whereas those attending Multi-Program Anatolian High Schools had an average of 391.14. As material shortages in schools are addressed, reading performance is expected to improve. Additionally, a shift in school location from village to city is predicted to increase a student's reading performance by approximately 14.60%. Reading scores also rise in tandem with socio-economic status, and the gender variable shows a positive difference in favor of female students. Lastly, participation in pre-school education was observed to contribute positively to students' reading skills. Moreover, Okatan (2021: 332) highlights in her literature review that reading comprehension is not only linked to verbal subjects like Turkish and literature but is also correlated with success in other disciplines such as mathematics and science. Likewise, Benzer's (2019: 107) study suggests that none of the reading comprehension questions in 5th, 7th and 8th grade Turkish language coursebooks represent PISA level 6 reading skills proficiencies. This can be regarded as another reason of the incompetency. Complementing this, Tokay Gedikoğlu (2005: 74) argues that Turkey lacks an ideology-free education system, and that newly implemented educational policies by successive governments have significantly, and often negatively, impacted student achievement.

CONCLUSION

Knowledge, the state of knowing, and systematically organized science play a central role in today's highly competitive world. They are crucial not only in shaping the interaction between nations and cultures but also in enabling countries to assert their presence on the global stage. Since the time of Plato, sciences have been classified in order to render them intelligible and to guide societies toward the philosophical ideals of "the good" and "happiness." This process of classification continued with Aristotle and has diversified over time, evolving into the modern understanding of scientific domains. From antiquity to modern times, philosophers have believed that classifying knowledge in various ways would be the most effective path forward and have accordingly proposed such classifications. For instance, Plato, emphasizing the primacy of universals, defined only theoretical disciplines as sciences. His student Aristotle, by contrast, emphasized that particulars also carry intrinsic meaning and classified disciplines such as physics, chemistry, biology, and medicine as sciences, grounded in empirical investigation and experimentation. The medieval Turkish-Islamic philosopher al-Fârâbî, while retaining elements of Aristotle's classification, developed his own distinctive schema. He placed the *Science of Language*—a field Aristotle did not consider a science at all—at the top of his classification system. Immediately following it, he positioned *Logic*, which Aristotle had treated not as a science but as a methodological tool guiding other sciences. Meanwhile, the *Science of Mathematics*, regarded by Plato as the most valuable and foundational discipline, was placed third in al-Fârâbî's hierarchy. From this, it can be concluded that al-Fârâbî regarded the *Science of Language* as indispensable to any cultural formation. In his seminal work *Ihsâu'l Ulûm (The Classification of the Sciences)*, he extensively examined the structure and general rules of the *Science of Language*, underscoring the idea that the ability to express thought is a precondition for a culture's intellectual enrichment.

Conversely, when viewed from a different angle, the ability to interpret written expressions, to draw inferences from a given text, and to synthesize both explicit and implicit meanings from ideas presented together is just as important as the act of expressing thoughts. In this regard, analyzing the reading proficiency data of national students as measured by international assessments is essential for policymakers in planning effective interventions. Studies (Benzer, 2019; Okatan, 2021) and PISA test results (OECD, 2019) reveal that 15-year-old students in Turkey consistently score below the international average in reading comprehension in their native language. This underperformance may stem from general factors such as school type and socio-economic background, but more specifically, it can also be attributed to the content of the instructional materials provided by the current education system—materials that may fail to equip students with the necessary skills for interpretation, analysis, and synthesis. From a broader perspective, the perpetually changing structure of the education system itself can be said to undermine the effectiveness of learning activities. Indeed, an examination of the views of Plato, Aristotle, and al-Fârâbî on the sciences clearly reveals that their ideas are interrelated and that mutual influence among them is both evident and inevitable.

When it comes to the most recent times, the PISA 2022 Turkey Report (2023) reveals that Turkish students scored 456 points—an even lower result than in 2009 which was 464 and in 2018 which was 466—reveals further contextual concerns. According to the 2022 report, 28% of students feel lonely at school, and 26% feel excluded. Compared to the 2018 report, the 2022 data suggest that students' sense of belonging at school has decreased (PISA 2022 Results, 2023: 5). Furthermore, 18% of students reported feeling unsafe on their way to school, 13% in their classrooms, and 20% in areas such as cafeterias, hallways, and restrooms (PISA 2022 Results, 2023: 6). Students' feelings of safety and belonging directly impact their academic performance, as evidenced by the data. This can be inferred as the incompetence of the efforts which have been made so far.

It can be said that raising students to become individuals who succeed in life and contribute meaningfully to society and culture involves a chain of interconnected and continuously interacting factors. Students' achievements in each academic discipline are directly related to their reading comprehension skills. Their ability to understand what they read, in turn, is influenced by a multitude of variables: from the level of family engagement in their education, to whether students feel a sense of belonging and safety at school; from the stability of the education system, to whether the curriculum is designed to foster specific competencies; and even to the type and location of the school, as well as the country's overall economic conditions. Therefore, the improvement of students' academic and social success requires coordinated and substantial efforts from the state, educators, and families alike.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

None.

Statements of publication ethics

We hereby declare that the study has not unethical issues and that research and publication ethics have been observed carefully.

REFERENCES

- Afşar, T. (1997). *Düşünce tarihi* (2. Baskı). İnsancıl Yayınları.
- Aristotle. (2009). "Metaphysics", *The basic works of Aristotle*, R. McKeon (Ed.), Random House Publishing Group.
- Arslan, A. (2006). *İlkçağ felsefesi tarihi sofistlerden Platon'a* (1. Baskı, C. 2). İstanbul Bilgi Üniversitesi Yayınları.

- Benzer, A. (2019). Türkçe ders kitaplarının PISA okuma yeterlik düzeyleri ile imtihanı. *Okuma Yazma Eğitimi Araştırmaları*, 7(2), 96-109. <https://doi.org/10.35233/oyea.659740>
- Cevizci, A. (2016). *İlkçağ felsefesi* (10. Baskı). Say Yayınları.
- Denzin, N. K., Lincoln, Y. S., Giardina, M. D., & Cannella, G. S. (Ed.). (2024). *The SAGE handbook of qualitative research* (Sixth edition). SAGE.
- Erdoğan, E. (2009). Platon ve Aristoteles'in bilimlere ilişkin sınıflamaları. *FLSF Felsefe ve Sosyal Bilimler Dergisi*, 7, 137-162.
- Farabî. (2019). *İlimlerin sayımı*. Türkiye İş Bankası Kültür Yayınları.
- Gedikoğlu, T. (2005). Avrupa Birliği sürecinde Türk eğitim sistemi: Sorunlar ve çözüm önerileri. *Mersin Üniversitesi Eğitim Fakültesi Dergisi*, 1(1), 66-80.
- Keat, R., & Urry, J. (2001). *Bilim olarak sosyal teori* (Nilgün. Çelebi, Çev.; 2. Baskı: Kasım 2001). İmge Kitabevi.
- Korkut, Ş. (2015). Fârâbî'de dil ve düşüncenin gelişim aşamaları. *Kutadgubilig Felsefe-Bilim Araştırmaları*, 28, 25-48.
- MEB. (2022). *PISA 2022*. T.C. Millî Eğitim Bakanlığı Ölçme, Değerlendirme ve Sınav Hizmetleri Genel Müdürlüğü.
- Mengüşoğlu, T. (2015). *İnsan felsefesi*. Doğu Batı Yayınları.
- OECD. (2010). *PISA 2009 Results: What Students Know and Can Do: Student Performance in Reading, Mathematics and Science (Volume I)*. OECD. <https://doi.org/10.1787/9789264091450-en>
- OECD. (2019). *PISA 2018 Results (Volume I): What Students Know and Can Do*. OECD. <https://doi.org/10.1787/5f07c754-en>
- OECD. (2023). *PISA 2022 Results (Volume I): The State of Learning and Equity in Education*. OECD. <https://doi.org/10.1787/53f23881-en>
- Okatan, Ö. (2021). PISA 2018 Türkiye Okuma Başarısının Bazı Değişkenler Açısından İncelenmesi. *Mehmet Akif Ersoy Üniversitesi Eğitim Fakültesi Dergisi*, 60, 331-353. <https://doi.org/10.21764/maeuefd.825174>
- PISA 2022 Results (Türkiye)*. (2023). [Factsheets]. OECD.
- Platon. (2010). *Devlet* (S. Eyüboğlu & M. A. Cimcoz, Çev.; 19.baskı). Türkiye İş Bankası Kültür Yayınları.
- Reeve, C. D. C. (2009). "Introduction", *The basic works of Aristotle*, R. McKeon (Ed.), Random House Publishing Group.
- Ross, D. (Ed.). (2005). *Aristotle* (2nd ed.). Taylor & Francis Ltd.