NEF EFMED

Cilt 16 - Sayı 2 - Aralık 2022

Necatibey Eğitim Fakültesi

Elektronik Fen ve Matematik Eğitimi Dergisi

Necatibey Faculty of Education

Electronic Journal of Science and Mathematics Education

Volume : 16

Issue : 2

Date : December 2022

ISSN : 1307-6086

NEF-EFMED (NFE-EJSME)

ISSN: 1307-6086

Necatibey Eğitim Fakültesi Elektronik Fen ve Matematik Eğitimi Dergisi (Necatibey Faculty of Education Electronic Journal of Science and Mathematics Education) Internet üzerinden ücretsiz yayın yapan yılda bir cilt, en az her ciltte iki sayı olarak yayımlanan, hakemli ve online bir fen ve matematik eğitimi dergisidir. Hedef kitlesi fen ve matematik eğitimcileri, fen ve matematik eğitimi öğrencileri, öğretmenler ve eğitim sektörüne yönelik ürün ve hizmet üreten kişi ve kuruluşlardır. Dergide, bu hedef kitlenin yararlanabileceği nitelikteki bilimsel çalışmalar yayımlanır. Yayın dili İngilizcedir.

Necatibey Faculty of Education, Electronic Journal of Science and Mathematics Education is an international on-line, refereed science and mathematics education journal that is published at least two issues in a year. NFE-EJSME is firmly established as the authoritative voice in the world of science and mathematics education. It bridges the gap between research and practice, providing information, ideas and opinions. It serves as a medium for the publication of definitive research findings. Special emphasis is placed on applicable research relevant to educational practice, guided by educational realities in systems, schools, colleges and universities. The journal comprises peer-reviewed general articles, papers on innovations and developments, research reports.

All research articles in this journal have undergone rigorous peer review, based on initial editor screening and anonymized refereeing by at least two anonymous referees. NEF-EFMED is an open access journal which means all content freely available without any charge. We support the rights of users to "read, download, copy, distribute, print, search, or link to the full texts of these articles".

Editör Kurulu

Necatibey Faculty of Education Electronic Journal of Science and Mathematics Education
Necatibey Eğitim Fakültesi Elektronik Fen ve Matematik Eğitimi Dergisi
Volume 16 Issue 1 June 2022 / Cilt 16 Sayı 1 Haziran 2022

CONTENTS / İÇİNDEKİLER

Önsöz

Herkese Merhabalar,

On on altıncı yılımızın ikinci sayısında toplam yedi makale yer almaktadır.

Bu sayıda katkıda bulunan gerek yazarlarımıza gerekse hakemlerimize çalışmalarından dolayı teşekkür ederiz.

Saygılarımla.

Editör

Dr. Hülya GÜR

Preface

Greetings to everyone,

In this edition of our journal, we have a total of seven articles related to science and mathematics education.

Thanks to everyone for contributing and/or becoming the reviewer of our journal.

Editor

Dr. Hülya GÜR

Necatibey Eğitim Fakültesi Elektronik Fen ve Matematik Eğitimi Dergisi
Cilt 16, Sayı 2, Aralık 2022, sayfa 275-303. ISSN: 1307-6086
Necatibey Faculty of Education Electronic Journal of Science and Mathematics Education
Vol. 16, Issue 2, December 2022, pp. 275-303. ISSN: 1307-6086

Research Article

The Effect of Problem-Solving Strategies Education on Developing Mathematical Literacy*

Hasan Temel [1], Murat Altun [2]

[1] Balıkesir University, Necatibey Faculty of Education, Department of Mathematics Education, Balıkesir/Türkiye, hasan.temel@balikesir.edu.tr, https://orcid.org/0000-0003-4532-0529

[2] Bursa Uludag University, Faculty of Education, Department of Mathematics Education, Bursa/Türkiye, maltun@uludag.edu.tr, https://orcid.org/0000-0001-8853-8523

Received : 19.09.2022 Accepted : 05.11.2022

Doi: 10.17522/balikesirnef.1164584

Abstract – One of the most significant areas of mathematics education today is how to improve students' mathematical literacy success and level. In this context, the study aimed to examine the effect of problem-solving strategies education on mathematical literacy. The study included 42 students in the eighth grade of secondary school, and the experimental design was preferred in quantitative research methods. Five-week problem-solving strategies education was conducted with 21 students in the experimental group. The study data were obtained with the mathematical literacy test (MLT), which was developed to determine the mathematical literacy achievement and levels of the experimental and control groups before and after education. As a result of the findings, it was determined that problem-solving strategies in education had a significant effect on students' mathematical literacy achievement. Concordantly, suggestions have been made regarding the need to use problem-solving strategies to improve mathematical literacy.

Key words: mathematical literacy, mathematics education, PISA, problem-solving, problem-solving strategies.

Corresponding author: Hasan TEMEL, Balıkesir University, Necatibey Faculty of Education, Department of Mathematics Education, Balıkesir/Türkiye.

Introduction

The Program for International Student Assessment [PISA], which is carried out every three years, was last carried out in 2018. Close to 3,5 million students from over 90 different

---

* This study was produced from the doctoral dissertation prepared by the first author under the supervision of the second author

countries have encountered reading, science, and mathematical literacy problems in PISA since 2000. More than 600,000 students participated in PISA 2018, representing nearly 32 million students aged 15 from 79 countries (Organisation for Economic Co-operation and Development [OECD], 2019). PISA focuses on how much of the knowledge and skills students can use in real life (Reinikainen, 2012; Stacey & Turner, 2015). In this context, the education given at school affects the success of countries in PISA.

If the education given at school is used in life outside school, it will be successful (Jacobs, 1989). One of the main aims of the curriculum should be to equip students with the knowledge, skills, and abilities to overcome real-life problems. According to some studies, however, the teaching contents are disconnected from real life (Mudaly, 2011; Syeda, 2015; Yıldırım, 1996; Yılmaz & Altınkurt, 2011). To overcome this problem, curricula should be related to real life (Altun & Akkya, 2014).

Understanding mathematics, which includes abstract structures, can also be achieved by associating it with real life. Learning mathematics "in addition to acquiring basic concepts and skills also includes thinking about mathematics, comprehending problem-solving strategies, and realizing that mathematics is an effective tool in real life" (Ministry of National Education [MoNE], 2013, p. 1). Among the school's main objectives, mathematics is raising individuals who can understand mathematical concepts, use them in real life, develop mathematical literacy skills, explain their thoughts in problem-solving, and explain the relationship between objects using the meaning and language of mathematics (MoNE, 2018). Considering this information, we face the need to plan training that will ensure the consolidation of connections. This will be done by focusing on the connection between school mathematics and real life in mathematics education. In order to meet this need, individuals require to gain skills such as making connections between thoughts, reasoning, estimation, and problem-solving as well as developing calculation skills (Colwell & Enderson, 2016). Because students need to encounter events that require using the knowledge they learned in the lessons and to provide opportunities to defend their suggestions by offering solutions to these events in terms of the use of mathematical knowledge in real life (Altun & Bozkurt, 2017).

Mathematical literacy (ML), which is seen as the level of use of mathematical knowledge in real-life, is an influential criterion for revealing the relationship between school mathematics and real-life situations. PISA, which offer a perspective on reading skills and science literacy as well as ML, are an application that evaluates to what extent 15-year-old students who are at the end or middle of compulsory schooling have acquired the knowledge and skills necessary for participation in modern society (OECD, 2016). PISA focuses not only on the evaluation of the knowledge that students have produced but also on how they can transfer what they have learned to new situations they may encounter inside and outside the school. In this context, PISA provides a way for the participating countries to review their education policies and practices. By keeping a mirror to the education systems of the countries, PISA reports have the opportunity to accurately describe the strengths and weaknesses of the current education systems of the countries, increase public awareness, evaluate education policies, curricula, and teacher competencies (Breakspear, 2012; Carvalho et al., 2017).

Türkiye took part in PISA, which was originally conducted in 2000, for the first time in 2003. It is seen that the scores at each basic area level are below the OECD average. In a way, this situation has opened Türkiye's education programs to discussion. Discussions took place not only in Turkey, but also in many countries such as Australia, Israel, South Africa and Finland. It can be said that the results of mathematics literacy within the framework of PISA are effective in the revision of the education programs of many countries (Breakspear, 2012). In this framework, the Ministry of National Education has accelerated its program development studies by determining the weaknesses in education in line with the PISA results and considering the results of international practices such as PISA in order to strengthen these aspects (Uyangör & Övez, 2012). In 2005, the mathematics curriculum in Turkey was changed and it was emphasized that the updated curriculum should teach mathematical concepts and expressions by relating them to real life and that students should gain skills such as problem-solving (Şefik & Dost, 2016). The effect of this change continued in the 2013 and 2018 curricula, and it was among the aims of real-life and ML curricula in mathematics curricula.

According to the PISA 2015 results, it is seen that Singapore ranks first in the field of ML (OECD, 2016) and ranks second according to the 2018 PISA results (OECD, 2019). When we examine Singapore's achievements in ML in other years, it ranked second similarly in PISA 2012 and 2009 applications. A look at the mathematics curricula for the secret of Singapore's success shows that, in 1992, the main goal of the Singapore mathematics program was mathematical problem-solving. In the curriculum, which was overhauled twice in 2001 and 2007, mathematical problem-solving remained the main purpose of the curriculum (Kaur & Yeap, 2009). Since the 1990s, mathematical problem-solving has been seen as the heart of the Singapore mathematics curriculum. It aims to develop student's skills such as problem-solving and process skills based on problem-solving (Toh, 2021). In countries such as Hong Kong and the Netherlands, which are at the top level in PISA and TIMSS, it is seen that problem-solving is given a significant place in the mathematics curriculum (Anderson, 2009). Considering that problem-solving is at the core of the mathematics curriculum of countries that are at the top in ML, it is thought that there is a relationship between problem-solving and ML. In this context, it is thought that examining the effect of problem-solving strategy education aimed at increasing the level of ML will offer a solution to improve the ML levels of countries that cannot achieve the desired success in the field of ML.

ML is closely connected with many concepts discussed in mathematics education (Stacey, 2015). One of these concepts is problem-solving. Because ML is generally determined by the solution processes for the problems presented in the context (Stacey & Turner, 2015). ML is based on developing problem-solving skills rather than developing basic mathematical skills (Ülger, et al., 2020). Problem-solving is the primary goal of being mathematically literate (Pugalee, 1999). In this direction, planning which strategies will be used in the problems encountered, in other words, the effective use of problem-solving strategies can be seen as an significant criterion for success in ML. Because problem-solving strategies are also used in the processes of transferring the problems given in the context of real life to the mathematical world, obtaining results from the problems in the mathematical world using mathematics, and interpreting, evaluating, and applying the obtained results to real-life situations. Therefore, the development of ML can be achieved with problem-solving strategies.

When the literature on ML is examined, studies examining the causes of failure in ML (Altun & Akkaya, 2014; Güler, 2013), and studies comparing the ML achievement levels of countries (Dossey et al., 2008; Kılıç et al., 2012), examining the ML achievements studies (Breen et al., 2009; Lin & Tai, 2015; Sáenz, 2009; Yenilmez & Uysal, 2011) and studies examining situations such as self-efficacy, attitude, and motivation towards mathematical literacy (Dinçer et al., 2014; Hopfenbeck & Kjærnsli, 2016) are encountered. Most of the studies conducted within the framework of ML are aimed at determining ML's situation (Ülger et al., 2020).

Purpose and Importance of the Study

When data from PISA examinations is presented on countries' ML levels, it is clear that many countries have not attained the intended results. This situation brings with it many discussions about education, and it is thought that innovations that will increase the ML level of students should be included in the curriculum. Therefore, determining the arguments that will increase the level of ML will be significant in terms of the effectiveness of the mathematics education given to the students. Therefore, this study aimed to examine the effect of problem-solving strategy education on ML achievement.

Recently, studies on ML in mathematics education have risen remarkably and have become the focus of educational research (Altun & Bozkurt, 2017; Rizki & Priatna, 2019; Wedege, 2010). When the studies on ML are analysed, it is seen that the majority of the studies are limited to descriptive definitions to reveal the existing situation, and there are not enough studies to increase the success levels of ML or to produce solutions to problems (Ülger et al., 2020). It is thought that this study will contribute to the literature in terms of presenting a teaching method and solution proposal to increase ML levels.

The research problems and sub-problems of this study, which aimed to determine the effect of problem-solving strategies on ML, were determined as follows:

1. What is the effect of problem-solving strategy education on students' mathematical literacy achievement level?"

Sub-Problems of the Research

1.1. Is there a significant difference between the mathematical literacy pre-tests of the experimental group and the control group?

1.2. Is there a significant difference between the mathematical literacy post-tests of the experimental group and the control group?

1.3. Is there a significant difference between the mathematical literacy permanency tests of the experimental group and the control group?

Method

In this study, a part of the thesis, the main topic of which is the classification of problem-solving strategies according to process skills, is reported. In this direction, the data obtained from the mathematical literacy test (MLT) used in the quantitative dimension of the research are included.

In order to examine the effect of problem-solving strategies education on ML, the quantitative research method, which is considered the preferred approach to test a theory or approach, was adopted (Creswell, 2013). Quantitative research is based on a systematic and pre-planned pattern. Since it is not possible to randomly assign individuals to groups in most educational research, it is seen that it is extremely difficult to use the real experimental design in educational research (Cohen et al., 2005). Therefore, since students cannot be randomly assigned to groups, quasi-experimental design is generally used in educational research (Evrekli et al., 2011). In addition, it is suggested that researchers use the quasi-experimental method when the sample cannot be determined randomly (Marczyk et al., 2005). The quasi-experimental design was preferred because the study was educational research, the groups were determined from the classes where teaching was carried out, and it was not possible to randomly assign the students participating in the study to the groups. The following table presents the research design:

Table 1. The research design of the experimental dimension

Groups	Pre-Test Data Collection Tool	Implementation	Post-Test Data Collection Tool	Persistence-Test Data Collection Tool
Experimental Group	MLT	Problem-Solving Strategies Education	MLT	MLT
Control Group	MLT	X	MLT	MLT

Study Group

The study group consists of 42 eighth-grade secondary school students. The research was carried out with students at eighth-grade level, since problem-solving strategies developed throughout the upper grades and required high-level skills (Işık & Kar, 2011) and because eighth-grade students were more successful at problem-solving than in the lower grades. In addition, the participation of students from the eighth-grade level at the secondary school level in PISA can be seen as another reason why the study was conducted at the eighth-grade level.

The students in the study group were divided into two groups, each comprise of 21 students, and these groups were randomly assigned as the experimental and control groups. Information about the experimental and control groups is presented in Table 2:

Table 2. Descriptive of study groups

Groups	Total (f)	Gender			
		Female		Male	
		f	%	f	%
Experimental	21	11	52.4	10	47.6
Control	21	9	42.9	12	57.1

Data Collection Tool

The "Mathematics Literacy Test (MLT)" was used to determine the ML achievement levels of the participants. MLT consists of ML problems used to determine students' ML levels in PISA 2000, 2003, and 2012. During the development of MLT, ML problems that were opened to access after declassification in PISA applications were examined, and the problems were examined according to mathematical content (Quantity, Space and Shape, Change and

Relationships, Uncertainty and Data) and ML levels. MLT consisting of 24 problems was created, four at each level for six mathematical literacy levels defined by PISA and six ML problems for each content area.

Content validity, sometimes expressed as logical or rational validity, was examined in order to determine to what extent MLT represents the situation to be measured (Shuttleworth, 2009). Content validity is an indicator of whether the items in the measurement tool adequately represent the situation to be measured (Büyüköztürk, 2013). Content validity is obtained within the framework of expert opinions (McMillan & Schumacher, 2010). In this context, expert opinion was sought to demonstrate the content validity of MLT. Opinions of three field experts (1 Prof. Dr., 1 Assoc. Prof. Dr., and 1 Assistant Prof. Dr.) were taken to evaluate the suitability of the study for the 24 problems in MLT. In line with the expert opinions, it was decided that the problems in the MLT were suitable for the purpose of the study and for the eighth-grade level.

MLT was also analysed for reliability, which is seen as the stability and consistency of test scores (Johnson & Christensen, 2014), obtaining similar results by using the same criteria, and the measurement being free of random errors (Karasar, 2008). In order to determine the internal consistency coefficient of the MLT, KR-20 reliability was calculated in line with the data obtained from 108 eighth-grade students outside the study group (KR-20 α=.88). In addition, analysis of the test-retest reliability criteria for MLT was also performed (Pearson Correlation Coefficient = .94). In this respect, it can be said that MLT is a reliable and stable measurement tool. MLT was applied to the experimental and control groups as a pre-test before the application, and as a post-test and permanency test after the application.

Implementation Process

In order to reveal the effect of problem-solving strategies education on the level of ML, five-week problem-solving strategies education was conducted with the experimental group. Separate lesson plans were created for each strategy for problem-solving strategies education. The lesson plans, which were created by considering the four stages of George Polya's problem-solving process (Polya, 1973), are designed to be carried out within one lesson hour for the training of each strategy. The opinions of five experts (1 Prof. Dr., 2 Assoc. Prof. Dr., 1 Assistant Prof. Dr., and 1 Ph.D. Student) were consulted on the relevance and intelligibility of the instructions in the designed plans and the compatibility of the related problems with the strategy to be implemented. Corrections were made to the lesson plans in line with expert evaluations. Apart from the study group (Fraenkel & Wallen, 2006), a pilot study was conducted with 18 students before the main implementation, in order to observe the surprise developments that were overlooked during the implementation, and to be aware of the changes that may occur in the operation of the application, and to observe the steps of the application (Fraenkel & Wallen, 2006). In line with the data obtained as a result of the pilot study, the lesson plans prepared for each problem-solving strategy were finalized and put into practice.

Before the implementation phase, studies in the literature on problem-solving strategies were examined (Alibali et al., 2009; Altun & Arslan, 2006; Altun et al., 2007; Eisenmann et al., 2015; Ramnarain, 2014; Verschaffel et al., 1999; Yazgan, 2007; Yazgan & Bintaş, 2005). In most studies where problem-solving strategy education was given, it was seen that the training was carried out by dividing the students into groups of two or three. Based on such groupings in the studies on problem-solving strategies education and the results of Nahornick's (2014) study of the effect of group and individual activities on solving open-ended non-routine mathematical problems, the results of group activities are more effective. It was decided to divide the training into groups. The students in the study group were divided into groups of three, with high, medium, and low-level students in each group, classified as high, medium, and low level according to the Transition from Primary to Secondary Education Exam [TPSE] results. With the experimental group, the strategies of "Make a Systematic List", "Guess and check", "Draw a Diagram", "Look for a Pattern", "Use Variable", "Simplify the Problem", "Work Backwards", "Make a Table" and "Logical Reasoning" In line with the lesson plans consisting of four problems for each strategy, problem-solving strategies training was carried out for five weeks (10 lesson hours). An example of the lesson plans created for problem-solving strategy education is given in the appendix.

Determined by Yıkmış (1999) and used by Pilten (2008) in calculating the reliability of problem-solving strategies education; the formula for Application Reliability=Number of Behaviours Displayed/Number of Total Behaviours x100 was used. The mean reliability of problem-solving strategies education was calculated as 94.1.

Data Analysis

The data obtained from the MLT for answering the research problems were transferred to the Statistical Package for the Social Sciences (SPSS) program, and statistical analyses were carried out to determine the significant difference between the scores of the experimental and control groups from the pre-test, post-test, and permanency tests. In order to determine the significant difference between the ML pre-tests, post-tests, and permanency tests of the experimental group, which was given problem-solving strategies education, and the control group, where the traditional teaching method was applied, first of all, the normality of the data was examined. The reason for examining the normality of the data groups is to decide which of the parametric or non-parametric analysis methods will be applied for the analyses to be applied to detect the significant difference (Büyüköztürk, 2013, Can, 2014; Karasar, 2008). As a result of normality tests, the parametric analysis method was used for normally distributed data groups, and the non-parametric analysis method was used for data groups that did not show normal distribution.

The summary table regarding the analyses conducted to determine whether there is a significant difference between the pre-tests, post-tests, and permanency tests of the experimental and control groups in line with the data obtained from the MLT is presented in table 3:

Table 3. Data analysis table

	MLT	MLT Content Areas			
		Quantity	Space and Shape	Change and Relationships	Uncertainty and Data
Pre-Test	Independent Sample T-Test	Mann-Whitney U	Mann-Whitney U	Mann-Whitney U	Mann-Whitney U
Post-test	Independent Sample T-Test	Mann-Whitney U	Mann-Whitney U	Independent Sample T-Test	Mann-Whitney U
Permanency Test	Independent Sample T-Test	Mann-Whitney U	Independent Sample T-Test	Mann-Whitney U	Mann-Whitney U

Findings and Discussions

Findings related to the first sub-problem

Prior to problem-solving strategies education, MLT was used as a pre-test for both experimental and control groups. In this direction, the independent sample t-test, which is one of the parametric analysis methods, was used to determine whether there is a significant difference between the pre-test scores of the experimental and control groups. The findings of the t-test analysis are presented in table 4:

Table 4. T-test analysis results regarding MLT pre-tests of experimental and control groups

MLT Pre-test	N	M	SD	df	t	p
Experimental	21	9.76	4.49	40	.03	.97
Control	21	9.71	3.87			

$p>.05$

When the t-test analysis table is examined, it is seen that the significance value ($p=.97$) is greater than the significance level of .05. Therefore, it can be stated that there is no significant difference between the mathematics literacy scale pre-test scores of the experimental group students who were given problem-solving strategies education and the control group students to whom the traditional teaching method was applied ($t_{(40)}$=.03, p>.05). In line with this finding, it can be said that there is no significant difference between the ML levels of the experimental and control groups. In addition, the groups are equivalent in terms of ML levels. In addition, as seen in the t-test analysis results, it is clear that the MLT pre-test average of the experimental group students ($M_{Experimental}$=9.76) and the MLT pre-test average of the control group students ($M_{Control}$=9.71) are remarkably close to each other.

Mann-Whitney U test analysis, which is one of the non-parametric analysis methods, was used to determine whether there was a significant difference between the MLT pre-tests of the experimental and control groups in terms of content areas. The results of the Mann-Whitney U analysis are presented in table 5:

Table 5. Results of Mann-Whitney U tests in terms of content areas of problems according to MLT pre-tests of experimental and control groups

Content Areas	MLT Pre-Test	N	Mean of ranks	Sum of ranks	U	p
Quantity	Experiment	21	22.62	475.00	197.00	.54
	Control	21	20.38	428.00		
Space and Shape	Experiment	21	22.88	480.50	191.50	.45
	Control	21	20.12	422.50		
Change and Relationships	Experiment	21	20.86	438.00	207.00	.72
	Control	21	22.14	465.00		
Uncertainty and Data	Experiment	21	19.74	414.50	183.50	.34
	Control	21	23.26	488.50		

p>.05

When the Mann-Whitney U analysis results are examined, it can be said that there is no significant difference between the MLT pre-tests of the experimental and control groups in terms of each content area (p>.05).

The correct response rates of the experimental and control groups to the problems in MLT according to the content area are presented in table 6:

Table 6. Results for the correct answer rates of the experimental and control groups according to the content areas for the MLT pre-tests

MLT Pre-Test	Quantity	Space and Shape	Change and Relationships	Uncertainty and Data
Experimental	66%	35%	33%	37%
Control	64%	29%	36%	43%

According to the content areas of the MLT pre-test problems of the experimental and control groups, it is seen that the highest rate of correct answers in both the experimental and control groups is in the "Quantity" content area. The lowest correct answer rate in terms of content in the experimental group was in the "Change and Relationships" ($Experimental_{Change\ and\ Relationships}$=33%) and the "Space and Shape" ($Control_{Space\ and\ Shape}$=29%) content area in the control group.

Findings related to the second sub-problem

The MLT post-test was administered to the experimental and control groups after problem-solving strategies education. Within the framework of the data obtained, it was concluded that the data groups showed a normal distribution. In this direction, an independent sample t-test was used to determine whether there was a significant difference between the MLT post-test scores of the experimental group and the control group. The independent sample t-test is presented in table 7:

Table 7. T-test analysis results regarding MLT post-tests of experimental and control groups

MLT Post-Test	N	M	SD	df	t	p
Experimental	21	13.04	4.66	40	2.06	.04*
Control	21	10.14	4.46			

*p<.05

According to table 7, it was concluded that there was a significant difference between the MLT post-tests of the experimental group who received problem-solving strategies education and the control group who continued their normal education ($t_{(40)}$=2.06, p<.05). When the MLT post-test mean scores are examined, it is seen that the post-test mean scores of the experimental group ($M_{Experimental}$=13.04) are higher than the mean scores of the control group ($M_{Control}$=10.14) ($M_{Experimental} > M_{Control}$). In line with these findings, it can be said that there is a significant difference in favor of the experimental group between the ML achievement levels of the experimental group students who were given problem-solving strategy education and the control group students to whom the traditional teaching method was applied.

In the analysis of the significant difference in terms of the content area of the problems according to the MLT post-tests, it was concluded that the "Change and Relationships" content area data showed a normal distribution, while the data groups related to the other content areas did not show a normal distribution. Therefore, an independent sample t-test was used to determine the significant difference between the MLT post-tests of the experimental and control groups according to the "Change and Relationships" content area, and the Mann-Whitney U test analysis was used to determine the significant difference according to the other content areas. In this direction, the results of the t-test and Mann-Whitney U analysis are included.

Table 8. T-test results in terms of the content area of changes and relationships according to MLT post-tests of experiment and control groups

Content Area	MLT Post-test	N	M	SD	df	t	p
Change and Relationships	Experimental	21	3,19	1.32	40	2.29	.02*
	Control	21	2,28	1.23			

*p<.05

Table 9. Results of Mann-Whitney U tests in terms of content areas of ML problems according to MLT post-tests of experiment and control groups

Content Areas	MLT Post-Test	N	Mean of ranks	Sum of ranks	U	P
Quantity	Experimental	21	21.31	447.50	216.50	.91
	Control	21	21.69	455.50		
Space and Shape	Experimental	21	25.95	545.00	127.00	.01*
	Control	21	17.05	358.00		
Uncertainty and Data	Experimental	21	26.62	559.00	113.00	.00*
	Control	21	16.38	344.00		

*p<.05

When the findings are examined in table 8 and table 9, it is seen that the significance value is less than .05, which is the significance level, compared to the other content areas except for the "Quantity" content area (p<.05). Therefore, it was determined that there was no difference between the MLT post-tests of the experimental and control groups according to the "Quantity" content area. In the areas of “Change and Relationships”, “Space and Shape”, and “Uncertainty and Data” content, it was concluded that there was a statistically significant difference in favor of the experimental group between the MLT post-tests of the experimental and control groups.

Table 10. Results for the correct answer rates of the experimental and control groups according to the content areas for the MLT post-tests

MLT Post-Test	Quantity	Space and Shape	Change and Relationships	Uncertainty and Data
Experimental	73%	44%	53%	60%
Control	69%	29%	38%	40%

When the correct answer rates of the students in the MLT post-test according to the content area are examined, it is seen that the highest rate of correct answers for both the experimental and control groups, as in the pre-test, is in the problems related to the "Quantity" content area. According to the MLT post-tests, while the correct answer rate of the problems in terms of the "Quantity" content area was close to each other in the experimental and control groups, it was observed that there were significant differences in the other content areas.

Findings related to the third sub-problem

The permanency test of the MLT was applied to both experimental and control groups in order to determine whether the problem-solving strategies education conducted with the experimental group continued to have an effect after it had ended. MLT permanency tests were conducted six weeks after the post-tests were applied to the experimental and control groups. Independent sample t-test was used to determine whether there was a significant difference between the MLT permanency test scores of the experimental and control group students.

Table 11. T-test results in terms of the content area of changes and relationships according to MLT permanency-tests of experiment and control groups

MLT Permanency Test	N	M	SD	df	t	p
Experimental	21	13.00	4.64	40	2.05	.04*
Control	21	10.23	4.04			

*p<.05

In line with the results obtained in table 11, it is seen that there is a significant difference between the MLT permanency tests of the experimental and the control group ($t_{(40)}=2.05$, $p<0.05$). According to the MLT permanency test mean scores of the experimental and control groups, the mean score of the experimental group ($M_{Experimental}=13.00$) is higher than the mean score of the control group ($M_{Control}=10.23$) ($M_{Experimental} > M_{Control}$). Therefore, it can be said that there is a significant difference in favor of the experimental group between the ML levels of the experimental and the control group according to the permanency tests.

When we analysed the MLT permanency tests of the experimental and control groups in terms of content areas, it was seen that the data groups showed a normal distribution according to the "Space and Shape" content area, but the data groups did not show a normal distribution according to the other content areas. In this regard, independent sample t-tests were used for normally distributed data groups, while the Mann-Whitney U test was used for groups that did not show a normal distribution.

Table 12. Results of t-test in terms of space and shape content area of ML problems according to MLT permanency tests of experimental and control groups

Content Area	MLT Permanency - Test	N	M	SD	df	t	p
Space and Shape	Experimental	21	2.42	1.16	40	2.34	.02*
	Control	21	1.61	1.07			

*p<.05

Table 13. Results of Mann-Whitney U tests in terms of content areas of ML problems according to MLT permanency tests of experiment and control groups

Content Areas	MLT Post-Test	N	Mean of ranks	Sum of ranks	U	P
Quantity	Experimental	21	21.79	457.50	214.50	.87
	Control	21	21.21	445.50		
Change and Relationships	Experimental	21	24.29	510.00	162.00	.12
	Control	21	18.71	393.00		
Uncertainty and Data	Experimental	21	26.26	551.50	120.50	.00*
	Control	21	16.74	351.50		

*p<.05

According to table 12 and table 13, it is seen that while there is no significant difference in the content areas of "Space and Shape" and "Uncertainty and Data" (p>0.05), there is a significant difference in the content areas of "Quantity" and "Change and Relations" (*p<0.05).

Table 14. Results for the correct answer rates of the experimental and control groups according to the content areas for the MLT permanency tests

MLT Permanency Test	Quantity	Space and Shape	Change and Relations	Uncertainty and Data
Experimental	%72	%40	%50	%62
Control	%66	%27	%40	%45

When we examined table 14, it was concluded that the content area with the highest correct answer rate in both the experimental and control groups was "Quantity", and the least correct response rate was in the "Space and Shape" content area.

Discussion and Conclusions

It is seen that efforts to make students realize the importance of mathematics in real life are increasing around the world (Bolstad, 2021). In this context, one of the main purposes of the education given at school should be to enable students to realize mathematics in real life and to use the knowledge and skills they learned at school for the solution of real-life problems. Mathematical literacy, which can be defined as the transfer of mathematical knowledge and skills to real life and the mathematical interpretation of real-life situations, is significant in solving real-life problems (Kabael & Ata Baran, 2019). When the studies conducted within the framework of mathematical literacy are examined, it is stated that these studies are generally aimed at determining the situation and there are not enough studies on how to increase the success of mathematical literacy (Ülger et al, 2020). Based on this situation, the aim of this study was to see if problem-solving strategies education had a significant influence on mathematical literacy achievement. Participants in the experimental group received problem-solving strategies training, and the mathematical literacy levels of the experimental and control groups were assessed before and after the training.

According to the MLT pre-tests, when the correct response rates of the problems in terms of the content area were examined, it was concluded that the area with the highest correct response rate in both the experimental and control groups was "Quantity". While the content area with the highest success rate in PISA 2003, 2006, and 2009 is uncertainty (İlbağı, 2012), it is seen that the content area with the highest success rate in PISA 2012 is “Quantity” (Anıl et al., 2015). While this finding of the study differs according to PISA 2003, 2006, and 2009, it is parallel to PISA 2012. It was concluded that the content area with the lowest correct answer rate was "Space and Shape". This finding of the study is similar to the PISA 2003, 2006, 2009, and 2012, and it is seen that it overlaps with Köse's (2012) findings. As in the general PISA exams, in terms of Türkiye, the subject area where students fail the most is “Space and Shape”. Similarly, it is stated that the most unsuccessful field in TIMSS exams is geometry. (Fidan & Türnüklü, 2010). The reason for the low success in the "Space and Shape" content area, which brings the field of geometry to mind (Altun, 2014), may be due to the fact that geometry, which is identified with nature and daily life, cannot be thought of in real life (Delice & Sevimli, 2010). In addition, the abstract perception of geometry in general causes students to move away from geometry (Baki & Özpınar, 2007) and reduces the success of these students in the "Space and Shape" content area. This situation can be shown as another reason for the low success of students in the "Space and Shape" content area.

Based on the results gained from the MLT post-tests of the experimental and control groups, it is thought that problem-solving strategies education has an important effect on improving students' ML success level. Because problem-solving activities improve students' questioning and thinking skills (Novita et al., 2012), problem-based instruction will increase students' ML success levels (Wardono, et al., 2016). In this direction, it was concluded that the problems used in the problem-solving strategies education increased the ML levels of the students. It can be said that this finding of the study supports the findings of Wardono et al. (2016). In addition, the findings obtained from the study differ from the finding that Muyo's (2015) activity-based mathematics education program did not experience a significant change in terms of ML. The effect of the problem-solving strategy education revealed in the study on the level of ML achievement is thought to be a guide for how countries can improve students' ML skills. In this context, it can be said that the results of this study are a guide for future studies within the framework of the idea (Ülger et al., 2020) that studies should be carried out to increase ML achievement levels or to produce solutions to problems in studies carried out within the framework of ML.

In terms of the content area of the experimental and control groups, it was seen that the most significant difference between the two groups in the MLT post-tests was in the content area of "Uncertainty and Data". Based on this finding, it is thought that the problem-solving strategies education carried out with the experimental group has an important effect on solving the problems related to the "Uncertainty and Data" content area. When the scores of the countries that are at the top of PISA 2012 are examined in terms of content, it is seen that they have achieved quite high scores in the content area of "Uncertainty and Data" (OECD, 2014). This situation is also expressed in the studies of Wheater et al. (2014). In addition, it is striking that the Netherlands, Vietnam, and Australia, which achieved significant success in ML, achieved high scores in the "Uncertainty and Data" content area. In this context, it is thought that the applications and training for the "Uncertainty and Data" problems within the framework of mathematical literacy will contribute significantly to the success of the countries.

As a result of the analysis of the MLT permanency tests, it was concluded that there was a significant difference in favor of the experimental group between the MLT permanency tests of the experimental and the control groups. This is an indication that education on problem-solving strategies has a permanent effect. It is thought that the permanent effect of problem-solving strategies education also affects the level of ML achievement. ML is associated with students' capacity to solve problems, reason, and produce effective solutions in different situations in daily life (Özgen & Bindak, 2008). In this context, it can be said that problem-solving strategies have an undeniable effect on creating solutions to ML problems. The findings obtained from the MLT pre-test, post-test, and permanency tests support these thoughts.

Therefore, we can infer that problem-solving strategies education accounts for the significant difference between the experimental and control groups in terms of ML. Because problems are at the centre of mathematics education and teaching in the realization of mental activities in the classroom (Lampert, 2001). Teaching problem-solving strategies contributes to the development of non-routine problem-solving skills (Altun & Arslan, 2006; Altun & Memnun, 2008; Artut & Tarım, 2006; Çelebioğlu & Yazgan, 2009; Elia et al., 2009). Considering that ML problems can also be viewed as non-routine problems (Altun, 2014), it is a significant finding of this study that problem-solving strategies education increases students' ML success levels. In this context, training can be carried out by making use of problem-solving strategies to improve students' ML skills and levels.

Limitations and Suggestions

In this study, nine problem-solving strategies were studied, and the effect of these strategies on ML was examined. In future studies, studies on the effects of different problem-solving strategies other than the nine strategies on ML can be conducted, and studies on which strategy contributes more to ML can be put forward.

In the study, problem-solving strategies education was limited to eighth-grade students. The effect of problem-solving strategies education on the level of ML can be determined by conducting training at different grade levels, with larger samples and longer durations. The fact that the participant group consisted of eighth-grade students and that these students were preparing for high school entrance exams did not us allow to support the findings obtained through observations and interviews with the students. In the study, the determination of the effect on ML was limited to MLT. In future studies, to reveal the effect of problem-solving strategies education in more detail, interviews and observations with students can be carried out and a holistic result can be revealed by supporting the quantitative findings with qualitative analyses.

No study in the literature examines the effect of problem-solving strategies education on students' ML success. In this respect, the findings of the research are seen as significant in terms of revealing the effect of problem-solving strategies on ML. In this direction, it is thought that studies examining the effect of problem-solving strategies on ML should be increased. It has been concluded that problem-solving strategies have a permanent effect on achievements in areas such as "Space and Shape" and "Uncertainty and Data", which are expressed within the framework of ML. Research can be conducted to determine the causes of the permanent effect in these areas and the permanent effects in the other two areas.

In this study, problem-solving strategies education was carried out with group studies, and it is thought that group studies have a significant effect on problem-solving strategies education. In this context, the effect of individual problem-solving strategy education on students' ML success can also be examined. In addition, considering the effect of group work, program developers can design programs and develop activities suitable for group work.

Matematik Okuryazarlığını Geliştirmede Problem Çözme Stratejileri Eğitiminin Etkisi

Özet:

Günümüzde matematik eğitiminin en önemli uğraşı alanlarının başında öğrencilerin matematik okuryazarlık başarı ve düzeyinin nasıl arttırılacağı gelmektedir. Bu bağlamda gerçekleştirilen çalışmada problem çözme strateji eğitiminin matematik okuryazarlığına etkisinin incelenmesi amaçlanmıştır. Araştırmaya ortaokul sekizinci sınıfta öğrenim gören 42 öğrenci katılmış ve nicel araştırma yöntemlerinden deneysel desen tercih edilmiştir. Deney grubunda bulunan 21 öğrenci ile beş haftalık problem çözme stratejileri eğitimi gerçekleştirilmiştir. Çalışmanın verileri, eğitim öncesi ve sonrasında deney ve kontrol grubunun matematik okuryazarlık başarı ve düzeylerini belirlemek amacıyla geliştirilen matematik okuryazarlık testi (MOT) ile elde edilmiştir. Elde edilen bulgular sonucunda, problem çözme stratejileri eğitiminin öğrencilerin matematik okuryazarlık başarı düzeylerine anlamlı bir etkisinin olduğu belirlenmiştir. Bu bağlamda matematik okuryazarlığını geliştirmede problem çözme stratejilerinden faydalanılması gerektiğine yönelik önerilerde bulunulmuştur.

Anahtar kelimeler: matematik okuryazarlığı, matematik eğitimi, PISA, problem çözme, problem çözme stratejileri

References

Alibali, M. W., Phillips, K. M., & Fischer, A. D. (2009). Learning new problem-solving strategies leads to changes in problem representation. *Cognitive Development*, *24*(2), 89-101. https://doi.org/10.1016/j.cogdev.2008.12.005

Altun M., & Memnun D. S. (2008). Mathematics teacher trainees' skills and opinions on solving non-routine mathematical problems. *Journal of Theory and Practice in Education*, *4*(2), 213-238.

Altun, M. (2014). *Ortaokullarda (5, 6, 7 ve 8. sınıflarda) matematik öğretimi* [Teaching mathematics in secondary schools (5th, 6th, 7th and 8th grades)]. (10th ed.). Alfa Aktüel.

Altun, M., & Akkaya, R. (2014). Mathematics teachers' comments on PISA math questions and our country's students' low achievement levels. *Hacettepe Üniversitesi Eğitim Fakültesi Dergisi [Hacettepe University Journal of Education]*, *29*(1), 19-34.

Altun, M., & Arslan, Ç. (2006). İlköğretim öğrencilerinin problem çözme stratejilerini öğrenmeleri üzerine bir çalışma [A study on elementary school students' learning problem solving strategies]. *Uludağ Üniversitesi Eğitim Fakültesi Dergisi*, *19*(1), 1-21.

Altun, M., & Bozkurt, I. (2017). A new classification proposal for mathematical literacy problems. *Education and Science*, *42*(190), 171-188. http://dx.doi.org/10.15390/EB.2017.6916

Altun, M., Memnun, D. S., & Yazgan, Y. (2007). Primary school teacher trainees' skills and opinions on solving non-routine mathematical problems. *Elementary Education Online, 6*(1), 127-143.

Anderson, J. (2009). *Mathematics curriculum development and the role of problem solving. The Australian Curriculum Studies Association's 2009 Biennial Conference (ACSA Conference).* Retrieved from http://www.acsa.edu.au/pages/images/-Judy%20Anderson%20-%20Mathematics%20Curriculum%20Development.pdf

Anıl, D., Özkan, Y. Ö., & Demir, E. (2015). *PISA 2012 araştırması ulusal nihai rapor* [PISA 2012 survey national final report]. Retrieved from https://drive.google.com/file/d/0B2wxMX5xMcnhaGtnV2x6YWsyY2c/view?resourcekey=0-eKJtL9If8HV5aaL2vclhMg

Artut, P. D., & Tarım, K. (2006). İlköğretim öğrencilerinin rutin olmayan sözel problemleri çözme düzeylerinin çözüm stratejilerinin ve hata türlerinin incelenmesi [Investigation of the elemantary school students' problem-solving levels, problem solving strategies and error types in the nonroutine word problems]. *Çukurova Üniversitesi Sosyal Bilimler Enstitüsü Dergisi, 15*(2), 39-50.

Baki, A. T., & Özpınar, İ. T. (2007). Logo destekli geometri öğretimi materyalinin öğrencilerin akademik başarılarına etkileri ve öğrencilerin uygulama ile ilgili görüşleri. *Çukurova Üniversitesi Eğitim Fakültesi Dergisi, 3*(34), 153.

Bolstad, O. H. (2021). Lower secondary students' encounters with mathematical literacy. *Mathematics Education Research Journal*, 1-17. https://doi.org/10.1007/s13394-021-00386-7

Breakspear, S. (2012). *The Policy Impact of PISA: An Exploration of the Normative Effects of International Benchmarking in School System Performance*. OECD Education Working Papers, No. 71, OECD Publishing, Paris, https://doi.org/10.1787/5k9fdfqffr28-en.

Breen, S., Cleary, J., & O'Shea, A. (2009). An investigation of the mathematical literacy of first year third-level students in the Republic of Ireland. *International Journal of Mathematical Education in Science and Technology*, *40*(2), 229-246. https://doi.org/10.1080/00207390802566915

Büyüköztürk, Ş. (2013). *Sosyal bilimler için veri analizi el kitabı* [Manual of data analysis for social sciences]. PegemA Yayıncılık.

Can, A. (2014). *SPSS ile bilimsel araştırma sürecinde nicel veri analizi* [Quantitative data analysis in the scientific research process with SPSS] (3rd ed.). Pegem Akademi.

Carvalho, L. M., Costa, E., & Gonçalves, C. (2017). Fifteen years looking at the mirror: On the presence of PISA in education policy processes (Portugal, 2000-2016). *European Journal of Education, 52*(2), 154-166. https://doi.org/10.1111/ejed.12210

Cohen, L., Manion, L., & Morrison, K. (2005). *Research methods in education* (5th ed.). Routledge Falmer.

Colwell, J., & Enderson, M. C. (2016). "When I hear literacy": Using pre-service teachers' perceptions of mathematical literacy to inform program changes in teacher education. *Teaching and Teacher Education, 53*, 63-74. https://doi.org/10.1016/j.tate.2015.11.001

Creswell, J. W. (2013). *Research design: Qualitative, quantitative, and mixed methods approaches*. Sage publications.

Çelebioğlu, B., & Yazgan, Y. (2009). İlköğretim öğrencilerinin bağıntı bulma ve sistematik liste yapma stratejilerini kullanma düzeyleri. *Uludağ Üniversitesi Eğitim Fakültesi Dergisi, 22*(1), 15-28.

Delice, A., & Sevimli, E. (2010). Geometri problemlerinin çözüm süreçlerinde görselleme becerilerinin incelenmesi: ek çizimler [Investigation of visualization ability in geometry problem solving process: auxiliary drawings]. *M.Ü. Atatürk Eğitim Fakültesi Eğitim Bilimleri Dergisi, 31*, 83-102.

Dinçer, B., Akarsu, E., & Yılmaz, S. (2016). İlköğretim matematik öğretmeni adaylarının matematik okuryazarlığı özyeterlik algıları ile matematik öğretimi yeterlik inanç düzeylerinin incelenmesi [The investigation of perceptions of math literacy self-efficacy and mathematics teaching efficacy belief levels of preservice primary mathematics teachers]. *Turkish Journal of Computer and Mathematics Education (TURCOMAT), 7*(1), 207-228. https://doi.org/10.16949/turcomat.99884

Dossey, J., McCrone, S., Turner, R., & Lindquist, M. (2008). PISA 2003 mathematical literacy and learning in the Americas. *Canadian Journal of Science, Mathematics and Technology Education, 8*(2), 140-152. https://doi.org/10.1080/14926150802169289

Eisenmann, P., Novotná, J., Přibyl, J., & Břehovský, J. (2015). The development of a culture of problem solving with secondary students through heuristic strategies. *Mathematics*

Education Research Journal, *27*(4), 535-562. https://doi.org/10.1007/s13394-015-0150-2

Elia, I., Van Den Heuvel-Panhuizen, M., & Kolovou, A. (2009). Exploring strategy use and strategy flexibility in non-routine problem solving by primary school high achievers in mathematics. *ZDM*, *41*(5), 605. https://doi.org/10.1007/s11858-009-0184-6

Evrekli, E., İnel, D., Deniş, H., & Balım, A. G. (2011). Fen eğitimi alanındaki lisansüstü tezlerdeki yöntemsel ve istatistiksel sorunlar [Methodological and statistical problems in graduate theses in the field of science education]. *İlköğretim Online,10*(1), 206-218.

Fidan, Y., & Türnüklü, E. (2010). İlköğretim 5. sınıf öğrencilerinin geometrik düşünme düzeylerinin bazı değişkenler açısından incelenmesi [Examination of 5th grade students' levels of geometric thinking in terms of some variables]. *Pamukkale Üniversitesi Eğitim Fakültesi Dergisi*, *27*(27), 185-197.

Fraenkel, J. R. & Wallen, N. E. (2006). *How to design and evaluate research in education*. McGraw-Hill.

Güler, H. K. (2013). Türk öğrencilerin PISA'da karşılaştıkları güçlüklerin analizi [Analysis of Turkish students' difficulties encountered in PISA]. *Uludağ Üniversitesi Eğitim Fakültesi, 26*(2), 501-522.

Hopfenbeck, T. N., & Kjærnsli, M. (2016). Students' test motivation in PISA: The case of Norway. *The Curriculum Journal*, *27*(3), 406-422. https://doi.org/10.1080/09585176.2016.1156004

Işık, C., & Kar, T. (2011). İlköğretim 6, 7 ve 8. sınıf öğrencilerinin sayı algılama ve rutin olmayan problem çözme becerilerinin incelenmesi [Investigating of the number perception and nonroutine problem solving skills of 6th, 7th and 8th grade students]. *Ahi Evran Üniversitesi Eğitim Fakültesi Dergisi*, *12*(1), 57-72.

İlbağı, E. A. (2012). *PISA 2003 matematik okuryazarlığı soruları bağlamında 15 yaş grubu öğrencilerinin matematik okuryazarlığı ve tutumlarının incelenmesi* [Examination of the mathematical literacy and attitudes of the students aged 15 in terms of PISA 2003 mathematical literacy items]. [Unpublished doctoral thesis]. Atatürk University, Erzurum.

Jacobs, H. H. (1989). *Interdisciplinary curriculum: Design and implementation*. Association for Supervision and Curriculum Development.

Johnson, B., & Christensen, L. (2014) *Educational research: Quantitative, qualitative, and mixed approaches* (5th ed.). SAGE Publications.

Kabael, T. & Ata Baran, A. (2019). Ortaokul matematik öğretmeni adaylarının matematik okuryazarlığı performanslarının ve matematik okuryazarlığına ilişkin görüşlerinin incelenmesi [Investigation of prospective middle school mathematics teachers' performances in, and conceptions of mathematical literacy]. *Eskişehir Osmangazi Üniversitesi Türk Dünyası Uygulama ve Araştırma Merkezi Eğitim Dergisi, 4*(2), 51-67.

Karasar, N. (2008). *Bilimsel araştırma yöntemi* [Scientific research method]. Nobel Yayın Dağıtım.

Kaur, B. & Yeap, B. H. (2009). Mathematical problem solving in Singaporeschools. In Kaur, B.,Yeap, B.H. & Kapur, M. (Eds.). *Mathematical problem solving: Yearbook 2009* (pp. 3-13). Association of Mathematics Educationand World Scientific.

Kılıç, S., Çene, E., & Demir, İ. (2012). Türkiye'deki matematik başarısının öğrenme stratejileri açısından 8 ülkeyle karşılaştırılması [Comparison of learning strategies for mathematics achievement in Turkey with eight countries]. *Kuram ve Uygulamada Eğitim Bilimleri, 12*(4), 2585-2598.

Köse, M. (2012). *PISA 2003, 2006 ve 2009 Türkiye Uygulaması matematik ortak maddelerindeki başarıların incelenmesi* [Examining Academic Achivement in PISA 2003, 2006 and 2009 Turkey Implementation by Using Mathematic Common Items]. [Unpublished master thesis]. Hacettepe University, Ankara.

Lampert, M. (2001). *Teaching problems and the problems of teaching*. Yale University Press.

Lin, S. W., & Tai, W. C. (2015). Latent class analysis of students' mathematics learning strategies and the relationship between learning strategy and mathematical literacy. *Universal Journal of Educational Research, 3*(6), 390-395. Doi:10.13189/ujer.2015.030606

Marczyk, G., DeMatteo, D. & Festinger, D. (2005). *Essentials of research design and methodology*. John Wiley & Sons.

McMillan, J. H., & Schumacher, S. (2010). *Research in education: Evidence-based inquiry* (7th ed.). Pearson

Ministry of National Education [MoNE]. (2013). Ortaokul matematik dersi 5-8. sınıflar öğretim programı [Middle school math class 5th -8th classes curriculum]. Talim Terbiye Başkanlığı Yayınları.

Ministry of National Education [MoNE]. (2018). *Matematik dersi öğretim programı (İlkokul ve ortaokul 1, 2, 3, 4, 5, 6, 7 ve 8. sınıflar)* [Mathematics lesson curriculum (Primary and secondary school grades 1, 2, 3, 4, 5, 6, 7 and 8)]. Retrieved from http://mufredat.meb.gov.tr/Dosyalar/201813017165445-MATEMAT%C4%B0K%20%C3%96%C4%9ERET%C4%B0M%20PROGRAMI%202018v.pdf

Mudaly, R. (2011). Risking it: Entering uneven socio-scientific spaces in a life sciences classroom. *African Journal of Research in Mathematics, Science and Technology Education*, *15*(3), 27-40. https://doi.org/10.1080/10288457.2011.10740716

Muyo, M. (2015). *Prizren Eğitim Fakültesi öğrencilerinin matematik okuryazarlığı problemlerini çözme becerilerinin geliştirilmesi* [Improving mathematical literacy and problem-solving abilities of students in Prizren faculty of education]. [Unpublished doctoral thesis]. Balıkesir University, Balıkesir.

Nahornick, A. (2014). *The effect of group dynamics on high-school students' creativity and problem-solving strategies with investigative open-ended non-routine problems.* [Unpublished doctoral thesis]. Columbia University, Columbia.

Novita, R., Zulkardi, Z., & Hartono, Y. (2012). Exploring primary student's problem-solving ability by doing tasks like PISA's question. *Journal on Mathematics Education*, *3*(2), 133-150. Retrieved from https://files.eric.ed.gov/fulltext/EJ1078578.pdf

Organisation for Economic Co-operation and Development [OECD]. (2014). *PISA 2012 Results: What Students Know and Can Do – Student Performance in Mathematics, Reading and Science (Volume I, Revised edition, February 2014).* PISA, OECD Publishing. http://dx.doi.org/10.1787/9789264201118-en

Organisation for Economic Co-operation and Development [OECD]. (2019), *PISA 2018 assessment and analytical framework*. PISA, OECD Publishing. https://doi.org/10.1787/b25efab8-en.

Organisation for Economic Co-operation and Development [OECD]. (2016). *PISA 2015 results (Volume I): Excellence and equity in education.* OECD Publishing. http://dx.doi.org/10.1787/9789264266490-en

Özgen, K., & Bindak, R. (2008). Matematik okuryazarlığı öz-yeterlik ölçeğinin geliştirilmesi [The development of self-efficiacy scale for mathematics literacy]. *Kastamonu Eğitim Dergisi*, *16*(2), 517-528.

Pilten, P. (2008). *Üstbiliş stratejileri öğretiminin ilköğretim beşinci sınıf öğrencilerinin matematiksel muhakeme becerilerine etkisi* [The effect of metacognitive instruction on mathematical reasoning of fifth grade primary school students]. [Unpublished doctoral thesis]. Gazi University, Ankara.

Polya, G. (1973). *How to solve it* (2nd ed.). Princeton University Press.

Pugalee, D. K. (1999). Constructing a model of mathematical literacy. *The Clearing House, 73*(1), 19-22.

Ramnarain, U. (2014). Empowering educationally disadvantaged mathematics students through a strategies-based problem solving approach. *The Australian Educational Researcher, 41*(1), 43-57. https://doi.org/10.1007/s13384-013-0098-8

Reinikainen, P. (2012). Amazing PISA results in Finnish comprehensive schools. In H. Niemi, A. Toom, & A. Kallioniemi (Eds.), *Miracle of education: The principles and practices of teaching and learning in Finnish schools* (pp. 3-18). Sense Publishers.

Rizki, L. M., & Priatna, N. (2019). Mathematical literacy as the 21st century skill. *Journal of Physics: Conference Series*, 1157(4), 1-5. https://doi.org/10.1088/1742-6596/1157/4/042088

Sáenz, C. (2009). The role of contextual, conceptual and procedural knowledge in activating mathematical competencies (PISA). *Educational Studies in Mathematics, 71*(2), 123-143. https://doi.org/10.1007/s10649-008-9167-8

Shuttleworth, M. (2009). *Content validity*. Retrieved from https://explorable.com/content-validity

Stacey, K. (2015). The real world and the mathematical world. In K. Stacey & R. Turner (Eds.), *Assessing mathematical literacy. The PISA experience* (pp. 57–84). Springer.

Stacey, K., & Turner, R. (2015). *Assessing mathematical literacy: The PISA experience*. Springer.

Syeda, T. J. (2015). Perceptions of Private College Teachers of Karachi about the Curriculum Prescribed by Sindh Bureau of Curriculum (BOC). *Acta Didactica Napocensia*, *8*(4), 13-24. Retrieved from https://files.eric.ed.gov/fulltext/EJ1092636.pdf

Şefik, Ö. & Dost, Ş. (2016). Secondary preservice mathematics teachers' views on mathematical literacy. *Necatibey Faculty of Education Electronic Journal of Science and Mathematics Education*, *10*(2), 320-338. https://doi.org/10.17522/balikesirnef.277935

Toh, T. L. (2021). School calculus curriculum and the Singapore mathematics curriculum framework. *ZDM–Mathematics Education*, *53*(3), 535-547. https://doi.org/10.1007/s11858-021-01225-6

Uyangör, S. M., & Övez, F. T. D. (2012). İlköğretim altıncı sınıf matematik dersi öğretim programı cebir öğrenme alanı kazanımlarına ulaşılma düzeyi [Reaching of attainment levels of algebra learning domain in sixth grade of elementary mathematics education program]. *Necatibey Eğitim Fakültesi Elektronik Fen ve Matematik Eğitimi Dergisi, 6*(1), 01-22.

Ülger, T. K., Bozkurt, I., & Altun, M. (2020). Thematic Analysis of Articles Focusing on Mathematical Literacy in Mathematics Teaching-Learning Process. *Education and Science*, *45*(201), 1-37. http://dx.doi.org/10.15390/EB.2020.8028

Verschaffel, L., De Corte, E., Lasure, S., Van Vaerenbergh, G., Bogaerts, H., & Ratinckx, E. (1999). Learning to solve mathematical application problems: A design experiment with fifth graders. *Mathematical Thinking and Learning, 1*(3), 195-229. https://doi.org/10.1207/s15327833mtl0103_2

Wardono, Waluya, S. B., Scolastika, M., & Candra, D. (2016). Mathematics literacy on problem based learning with indonesian realistic mathematics education approach assisted e-learning edmodo. In *Journal of Physics: Conference Series*, 693(1), 1-10. IOP Publishing.

Wedege, T. (2010). Ethnomathematics and mathematical literacy: People knowing mathematics in society. In C. Bergsten, E. Jablonka, & T. Wedege (Eds.), *Mathematics and mathematics education: Cultural and social dimensions. Proceedings of MADIF 7*.

The Seventh Mathematics Education Research Seminar, Stockholm (pp. 31–46). Linköping: Linköping Universitet.

Wheater, R., R. Ager, B. Burge, & J. Sizmur. (2014). *Achievement of 15-year-olds in England: PISA 2012 national report (OECD Programme for International Student Assessment).* Retrieved from https://dera.ioe.ac.uk/19484/1/programme-for-international-student-assessment-pisa-2012-national-report-for-england.pdf

Yazgan, Y. (2007). Dördüncü ve beşinci sınıf öğrencilerinin rutin olmayan problem çözme stratejileriyle ilgili gözlemler [Observations about fourth and fifth grade students' strategies to solve non-routine problems]. *İlköğretim Online, 6*(2), 249-263.

Yazgan, Y., & Bintaş, J. (2005). İlköğretim dördüncü ve beşinci sınıf öğrencilerinin problem çözme stratejilerini kullanabilme düzeyleri: Bir öğretim deneyi. *Hacettepe Üniversitesi Eğitim Fakültesi Dergisi, 28*, 210-218.

Yenilmez, K., & Uysal, E. (2011). Sekizinci sınıf öğrencilerinin matematik okuryazarlığı düzeyi [The mathematics literacy level of eighth grade students]. *Eskişehir Osmangazi Üniversitesi Sosyal Bilimler Dergisi, 12(*2), 1-15.

Yıkmış, A. (1999). *Zihin engelli çocuklara temel toplama ve çıkarma işlemlerinin kazandırılmasında etkileşim ünitesi ile sunulan bireyselleştirilmiş öğretim materyalinin etkililiği* [The effectiveness of the individualized teaching materials presented with interactive unit on the teaching of the basic addition and subtraction skills to the student with mental retardation]. [Unpublished doctoral thesis]. Anadolu University, Eskişehir.

Yıldırım, A. (1996). Disiplinlerarası öğretim kavramı ve programlar açısından doğurduğu sonuçlar. *Hacettepe Üniversitesi Eğitim Fakültesi Dergisi,* 12, 89-94.

Yılmaz, K., & Altınkurt, Y. (2011). Öğretmen adaylarının Türk eğitim sisteminin sorunlarına ilişkin görüşleri [Prospective teachers' views about the problems of Turkish educational system]. *Uluslararası insan bilimleri dergisi, 8*(1), 942-973.

Appendix

Lesson plan : Lesson 1
Subject : Make a Systematic List Strategy
Duration : 40 min
Group : 3-4 student
Tools : Paper, pencil, Worksheets, Projector
Behaviours : 1- Understanding the Problem
2- Ability to Think About the Problem
3- Exploring the Make a Systematic List Strategy
4- Implementing the Make a Systematic List Strategy
5- Naming Problems Related to the Make a Systematic List Strategy

	Learning Activities	**Behaviours Expected from Students**	**Guiding Behaviours**
Understand the problem	If two coins are tossed at the same time, what is the number of possible pairs?	Determining what is given and desired within groups	Students are given 5 minutes to discuss the problem. What is given in the problem? What is required? How are the front and back of a coin named?
Devise a plan	First coin: Heads Second coin: Heads or Tails First coin: Text	Students are expected to argue among themselves for a solution. Students are expected to decide to make a systematic list as a result of their discussion.	What can be the values of the first coin tossed on the upper surface? What can be the values on the upper surface of the second coin?
Carry out the plan	Second coin: Heads or tails HH, HT, TH, TT	The determining strategy is implemented. Students are expected to write down what values the second coin can get when the first coin systematically comes "heads". When the first coin comes "Tails", they are expected to write down what values the second coin can take.	Meanwhile, the teacher goes around the groups and checks the students' practices.
Look back	Four pairs	Students are expected to discuss their results among themselves and evaluate the result.	Groups are asked to describe their results on the board.

	Learning Activities		**Behaviours Expected from Students**	**Guiding Behaviours**
Understand the problem	Below is a restaurant's menu. <u>Pre Meals</u> Tomato soup Stuffed with Olive Oil	 <u>Main Meals</u> Steak Meaty Chickpeas Mushroom Sauteed	Determining what is given and desired within groups	What is given and what is required in the problem? How many of the pre-meal and main meals should be chosen? Can more than one be chosen from the pre-meal?

	How many different ways can you eat with this menu, provided that you choose one each from the pre-course and main course sections?		Can more than one be chosen from the main meal?
Devise a plan	Pre-Meal Tomato soup Main Meal Steak Meaty Chickpeas Mushroom Sauteed	Students are expected to argue among themselves for a solution. Students are expected to decide to make a systematic list as a result of their discussion.	What can be chosen as a pre-meal? What can be chosen as the main meal?
Carry out the plan	Three situations Pre-meal: Stuffed with Olive Oil Main Meal Steak Meaty Chickpeas Mushroom Sauteed Three situations	Students are expected to write down the pre-meal and main meals systematically.	Meanwhile, the teacher goes around the groups and checks the students' practices. For struggling groups, what can be the main course when pre-meal tomato soup is chosen? What can be the main course when the pre-meal is "stuffed with olive oil"? leading questions can be asked.
Look back	Six different situations	Students are expected to discuss their results among themselves and evaluate the result.	Group spokespersons are invited to the board and asked to share the solutions they found. Ask students if they have different ideas.

	Learning Activities	Behaviours Expected from Students	Guiding Behaviours
Understand the problem	How many different two-digit numbers with different numerical values can be written using the numbers 2, 5, and 8?	Writing down what is given and what is wanted, The numbers 2, 5, and 8 are given. A two-digit number is required.	
Devise a plan	numbers starting with 2: 25, 28	Making a systematic list	It is asked whether there are similarities with other solved questions.
Carry out the plan	numbers starting with 5: 52, 58 numbers starting with 8: 82, 85	The smallest number that can be written is 25, The largest number is 85. All possible cases should be listed. Two-digit numbers starting with 2, 5, and 8 are determined.	The teacher checks the students by walking around the groups and poses guiding questions to the groups that cannot find the solution.
Look back	6 different numbers	The question is reviewed within the group. It is debatable why numbers such as 22, 55, and 88 are not written.	It is emphasized that the critical point in such problems is to determine where to start the sequence and to act systematically. It may be asked why 22 was not written.

	Learning Activities	Behaviours Expected from Students	Guiding Behaviours
Understand the problem	How many kinds of 25 liras can you get with the 1, 5, 10 liras you have?	There are 1, 5, and 10 liras. It is requested to obtain 25 liras. By using 1, 5, and 10 liras, 25 liras will be obtained.	What are the givens? What is desired?
Devise a plan	25 = 10 + 10 + 5 10 + 10 + 1 + 1 + 1 + 1 +1 10 + 5 + 5 + 1 + 1 + 1 + 1 +1 10 + 5 + 1 + 1 + 1 + 1 +1 + 1 + 1 + 1 + 1 +1 10 + 1 + 1 + 1 + 1 +1 + 1 + 1 + 1 + 1 +1 + 1 + 1 + 1 +1+ 1	A list must be made. In the list to be made, the use of 1, 5, and 10 liras should be determined systematically. Making a Systematic List	Have you encountered such problems before?
Carry out the plan	10 + 5 + 5 + 5 5 + 5 + 5 + 5 + 5 5 + 5 + 5 + 5 + 1 + 1 + 1 + 1 +1 5 + 5 + 5 + 1 + 1 + 1 + 1 +1+ 1 + 1 + 1 + 1 +1 5 + 5 + 1 + 1 + 1 + 1 +1+ 1 + 1 + 1 + 1 +1+ 1+ 1+ 1 + 1 +1 5 + 1 + 1 + 1 + 1 +1+ 1 + 1 + 1 + 1 +1+ 1 + 1 + 1 + 1 +1 + 1 + 1 + 1 + 1 +1 1 + 1 + 1 + 1 +1+ 1 + 1 + 1 + 1 +1+ 1 + 1 + 1 + 1 +1 + 1 + 1 + 1 + 1 +1+ 1 + 1 + 1 + 1 +1	Using 10, 5, and 1 lira, the production of 25 liras is written systematically. The cases where the 10 liras are used twice and the cases where they are used once are written. The cases where 5 liras are used 5 times, 4 times, 3 times, 2 times and 1 time are written. The situation in which 25 liras is obtained by using only 1 lira is determined.	The teacher checks the students by walking around the groups and poses guiding questions to the groups that cannot find the solution.
Look back	12 Situation	In such a case, the most important point is to determine where to start the ranking.	The lists created by the groups are written on the board. The list of each group is examined separately, and it is questioned how they created this list. The groups that make up the list are asked to explain according to which system they created the list. At the end of the lesson, the groups are asked to name the strategy used to solve the problems, and the strategy used throughout the lesson is named in line with the common decision of the class.

Necatibey Eğitim Fakültesi Elektronik Fen ve Matematik Eğitimi Dergisi
Cilt 16, Sayı 2, Aralık 2022, sayfa 304-325. ISSN: 1307-6086
Necatibey Faculty of Education Electronic Journal of Science and Mathematics Education
Vol. 16, Issue 2, December 2022, pp. 304-325. ISSN: 1307-6086

Research Article

Secondary school mathematics teachers' understanding of materials and materials they create

Emine Nur ÜNVEREN BİLGİÇ[1], Ziya ARGÜN[2]

[1] Duzce University, Düzce-Türkiye, eunveren@sakarya.edu.tr, http://orcid.org/0000-0001-9684-4192

[2] Gazi University, Ankara-Türkiye, ziya@gazi.edu.tr, http://orcid.org/0000-0001-8101-7215

Received : 27.06.2022 Accepted : 12.11.2022
Doi: 10.17522/balikesirnef.951308

Abstract – The study aimed to examine mathematics teachers' materials usage in instructional design and was carried out via qualitative paradigm of action research with eight teachers working at the Ministry of National Education. Eleven sessions were held in total and data were gathered through semi-structured interviews, written documents, video recordings and focus group interviews. The data were analyzed via content analysis and constant comparative analysis. The results revealed while manipulatives in pre-education could be only observed by students; in post-education they were experienced by students, with teacher acting as a guide.

Key words: teacher training, instructional design, manipulative, virtual material.

Corresponding author: Emine Nur ÜNVEREN BİLGİÇ, eunveren@sakarya.edu.tr

Introduction

In today's world, students should be able to collect necessary information, model, solve problems and moreover foresee possible mathematical situations when they encounter real-life problems. In this case, students are expected to construct the process during mathematics teaching. Teacher should only be the guiding individual. On the other hand, in this student-centered approach students end their problematic mathematical exercises they started themselves with a related mathematical situation through their own actions and efforts. Using information and communication technologies appropriately and efficiently is of great importance in this process and it is one of the components which will complete and facilitate this program's implication. Therefore, teachers should bring to the class by planning well-structured activities. (The Ministry of National Education, 2013).

Materials usage is supported by several theories in education (Dienes and Golding, 1971; Piaget, 1971; Skemp, 1987). Learning theorists suggested that individuals are greatly influenced by their environment while constructing concepts in their minds; therefore materials play a crucial role at this stage (Post, 1988). Piaget (1971) stated that individuals construct concepts not as a copy of reality but as a reconstruction. Dewey (1938) claimed that the primary requirement of teaching programs should be real life experiences. Bruner (1960) stated that learning is a process rather than a product; while Dienes (1969) emphasized that mathematics teachers should construct and create the concepts in students' minds rather than transferring them. Piaget suggested that primary school students are at concrete learning stage; therefore we need to appeal to their sense organs with a number of concrete materials (İnan, 2013). Piaget (1971) claimed that mathematical information occurred in individuals through relating. This type of information is different from physical information which is gathered through observation of the outside world. According to Piaget, physical information can be constructed through observation or experiment; whereas mathematical information can be constructed through thought-based abstractions. Additionally, Bruner's (1966) learning theory suggests that in order for a meaningful learning to occur in an individual, there should be teaching based on individual's exploration and discovery of new relationships. Educational materials boost students' interest in lessons, facilitate learning and boost their motivation (Yalın, 2000).

Whatever the theoretical benefits of teaching materials are, using them requires a specific level of knowledge and skill. In this respect, teacher who has an important place in the teaching system organizes, manages and supervises other components in the teaching system, determined the tools to be used in teaching, applies teaching techniques and comments on the results (Abrami, 2001). If the teacher does not possess the skills to use these materials, s/he will not want to use them. Besides, whether or not using materials is not related to only knowledge and skills. The teachers need to believe their benefits and be willing to use them (Holmes, 2013; Haara, 2010). On the other hand, Domino (2010) states that if teachers receive training on material usage, they will be influenced to use materials in their own lessons (Marshall and Swan, 2005; McClung, 1998). Again, according to Bozkurt and Akalın (2010) and Moyer (2001) materials are tools which are used to make abstract concepts concrete and make teaching more effective in teaching environments (Cope, 2015; Hartshorn & Boren, 1990; Laski, Jor'dan, Daoust, & Murray, 2015; McClung, 1998; Moyer, 2001; Ojose & Sexton, 2009; White, 2012, Holmes, 2013). Materials are used in order to materialize abstract concepts and teach more effectively (Boggan, Harper & Whitmire, 2010; Cope, 2015; Hartshorn & Boren, 1990; Laski, Jor'dan, Daoust & Murray, 2015; McClung, 1998; Moyer, 2001; Ojose & Sexton, 2009; White, 2012).

The present study also asked teachers' opinions on materials and examined materials that they used in their instructional design after a related education as well as in pre-education. On this subject, Domino (2010) stated that teachers' education would influence their materials usage in their classes. Therefore, this study is quite important for materials usage in teaching in the sense that it examines the influence of teachers' education on their teaching material design or views. It is expected to inspire researchers who will study on this field. Additionally, results of the study are expected to contribute greatly in teacher education and in-service teacher training.

So, an answer to the question of "How is the teachers' perceptions and (if there is any) experiences regarding material usage before and after materials design training in the process of teaching design?" was sought. In this context, the following sub-problems were answered:

(1) What are the perceptions of the teachers regarding materials design before the training?

(2) Did the teachers use materials in the process of teaching design before receiving training? If so, what kind of materials did they use?

(3) What are the teachers' perceptions regarding materials usage in the process of teaching design after receiving training?

(4) Do the teachers use materials in the process of teaching design after the training? If so, what kind of materials do they use?

Method

The study began with the researchers defining the problem situation by examining the literature. Then, an interview was done with the head of mathematics teachers' group in the most successful school in the present district about the problem case (Appendix, A). Content of the semi-structured interview was analyzed via content analysis. The analysis revealed 4 themes which are need for materials, materials' place and importance in teaching environment, types of materials and learning outcomes of materials. Learning outcomes at the end of the interviews were defined as explaining similar triangles, irrational numbers and real numbers; showing and explaining intercept of triangle medians, defining the relationship between roots and coefficients of second-degree equation with one unknown, calculating the number of different ways to choose r elements from n-element set; displaying and using features regarding modular arithmetic; interpreting distance using minimized or maximized drawing of an object;

interpreting an object's perimeter, area or volume; explaining a function's limit thorough examples, explaining the concept of derivatives by using physics and geometry models and guessing the area of limited region between a function's graphic and x-orbit through Riemann sum. According to these learning outcomes, lesson plans were developed with the contributions of four mathematics educators including the researchers. After this interview, problem case was defined clearly; 10 learning outcomes were defined within the framework of Mathematics Teaching Program; and an action plan was developed with the help of the interviewee teacher. After this step, the participants of the study were defined. To define the participants, firstly all the names of the in-service mathematics teachers were received from the District Directorate of National Education. Then all the teachers were informed about the study and asked whether they would like to participate in it. An introductory meeting was held with the twelve teachers who agreed to take part in the study. At the end of the introductory meeting, three teachers stated that they would not be able to participate because of private reasons. After the training sessions started, one of the teachers could not participate because of health problems and the final participant number remained as eight. In choosing the participants, criterion sampling was used as the training took ten weeks, the study was video-recorded and the participants were expected to display products.

The researcher held sessions on Thursdays between 17.00- 18.15 with the participants. These sessions were held in the computer laboratory in B Block room 2205 at X University, Faculty of Education. First session included semi-structured interviews with the participants (Appendix B). Next, the researcher informed the participants on "A Closer Look at the Concept of Materials" for two hours. Following sessions included ten weeks of training for ten learning outcomes. At the beginning of these training sessions, the participants were asked to hand in their relevant instructional designs that they shared with their students. The researcher gave training regarding the relevant learning outcome at the same session. At the end of the session, the participants were asked to design a teaching in accordance with the training they received. The participants were asked to hand in the second instructional design in next week's session (Certain participants asked to for longer time to hand in their designs). Semi-structured interviews were done with certain participants in order to acquire thorough information regarding their instructional designs.

Considering the research process, the study is liberating/constructive/critical action research which is one of action research types. The critical study (Norton, 2009) aims for the appliers gaining new information, skills and experiences as well as a critical point of view toward their own implications. Thus, the applier will consider their implications as a problem-solving process and examine their role in this process. Also, they will acquire the ability to review their implications critically and explain the frequently encountered problems rationally (Yıldırım & Şimşek, 2008).

Data collection

Multiple data gathering tools were used in the study. The reason for this was the effort to see through a supportive and integrated perspective. The data gathering tools which were used in the study were; semi-structured interviews, focus group interviews, written documents and video recordings.

The interviews were carried out with an interview form approach. Questions on the interview form were prepared by the researcher with reference to the literature (Kay & Knack, 2008; Domino, 2010; Marshall & Swan, 2005; Kamii, Lewis & Kirkland, 2001). The forms were examined by three different experts. The interviews forms were presented in the appendix. Intention of the first interviews before the training session was to define teachers' views on materials usage. Therefore, the teachers were asked eleven questions. The data gathered from these interviews were repeatedly read, coded and divided into themes by the researcher (Appendix B).

The researcher defined ten instructional designs related to these learning outcomes in accordance with the instructor's and mathematics teacher's opinions. The learning outcomes were based on Article 13 which stated the necessary points to consider while preparing a lesson plan in Instruction regarding Planned Practice of Education of the Ministry of National Education (2005). The participants were given the concrete manipulative materials below which were developed by prospective teachers within the scope of Special Teaching Methods class. In addition, the researcher prepared a number of worksheets using GeoGebra Software.

Figure1. Concrete manipulatives used in the study (wooden skewers, mode board, suction hook triangle plate).

All of the training in the study was carried out as a focus group interview. The interviews were video recorded. The focus group interviews used 5-question interviews form. However, beside the questions on the form there were attempts to go deeper with the occasional drill questions. The interview form is presented in the appendix (Appendix C).

Another data gathering tool on the study was the instructional designs of the teachers. These instructional designs were divided into two stages. These stages were 1) pre-training instructional designs and 2) post-training instructional designs.

One of the data gathering tools was video recording. These recordings included all the interviews and training. Consent of the participants was taken at the beginning of video recording. The recordings occurred in a way that would allow the participants and the researcher to be seen easily but not bother them.

Data Analysis

Video recordings gathered from the interviews and focus group interviews were transcribed. The researcher tried to reflect the pauses and moods of the participants (such as loud voice, laughing) as much as possible. The researcher presented these details in parentheses.

Written documents gathered from the interviews were put to content analysis by the researcher. Certain codes were defined and then these codes were transferred into themes. While defining codes and themes; open coding and axial coding, which are types of coding techniques of grounded theory.

Written document gathered from the teachers (instructional designs) were put to document analysis. The researcher digitized this data set according to certain categories from other data resources. Digitization process used "existent" or "non-existent" system. Thus, if the relevant category is present in a document, the value would be 1; while if no present, it would be 0 (Yıldırım & Şimşek, 2006). The categories of document analysis of the teachers' instructional designs were as follows:

(1) Is there concrete material in the instructional design?

(2) Is there virtual manipulative in the instructional design?

(3) Are the steps clear in the instructional design?

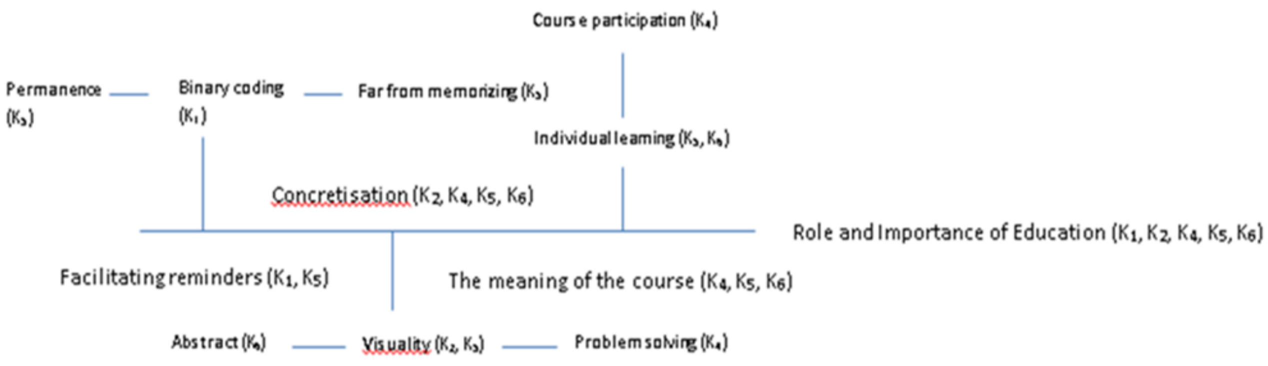
(4) Is there a Plan B in the instructional design (in case the students cannot learn)?

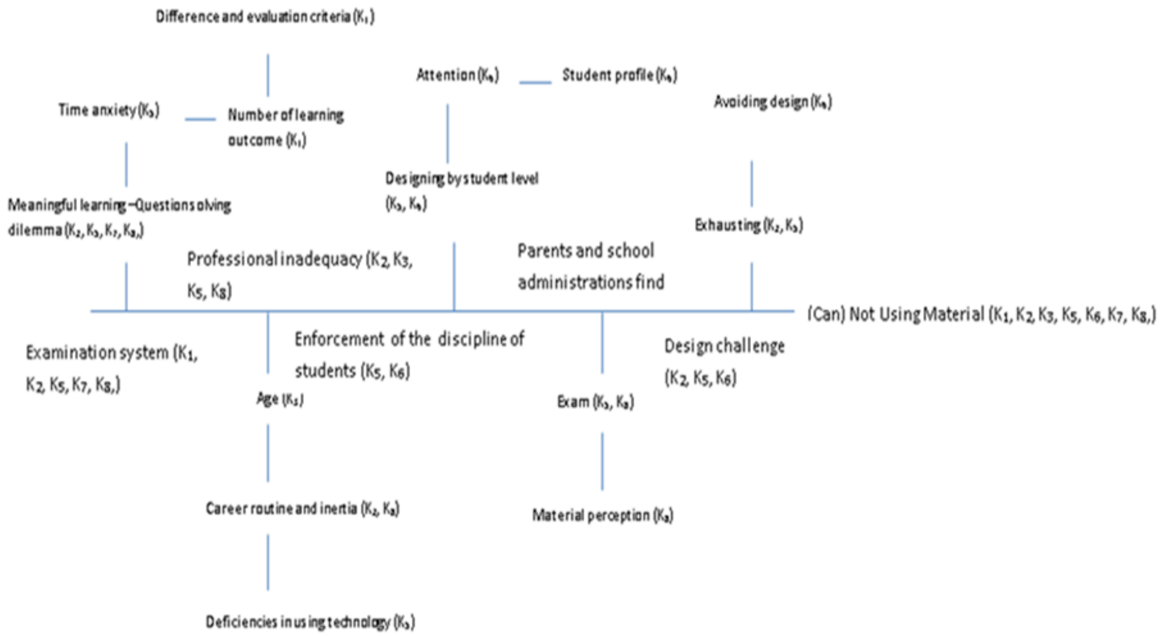
Validity and reliability

Persuasiveness in the present study was contacting the participants for ten weeks, various data (confirming the data via different data sources), and confirmation of interviews reports by the interviewees and evaluation of reports by two experts. The present study's transferability was elaborate explanation of the participants and study process and usage of purposeful sampling. The researcher gathered the data in similar processes. Additionally, consistency analysis was carried out through data coding and their relationships. With this purpose, the researcher described her role during the study. Additionally, data analysis process was explained elaborately and the data were coded twice.

Findings and Discussions

Five categories were defined in total from the interviews with the teachers regarding the study's first sub-problem. These categories were Material perception, Its place and importance in teaching, Not using/being able to use materials, The Place the material is used and The Priorities in Designing and Preparing Material. The axial coding regarding these categories and sub-categories are shared below:


Course participation (K_4)
Permanence (K_3)
Binary coding (K_1)
Far from memorizing (K_3)
Individual learning (K_3, K_6)
Concretisation (K_2, K_4, K_5, K_6)
Role and Importance of Education (K_1, K_2, K_4, K_5, K_6)
Facilitating reminders (K_1, K_5)
The meaning of the course (K_4, K_5, K_6)
Abstract (K_6)
Visuality (K_2, K_3)
Problem solving (K_4)


Difference and evaluation criteria (K_1)
Attention (K_6)
Student profile (K_6)
Time anxiety (K_3)
Number of learning outcome (K_1)
Avoiding design (K_6)
Designing by student level (K_3, K_6)
Meaningful learning –Questions olving dilemma (K_2, K_3, K_7, K_8,)
Exhausting (K_2, K_3)
Professional inadequacy (K_2, K_3, K_5, K_8)
Parents and school administrations find
(Can) Not Using Material (K_1, K_2, K_3, K_5, K_6, K_7, K_8,)
Examination system (K_1, K_2, K_5, K_7, K_8,)
Enforcement of the discipline of students (K_5, K_6)
Design challenge (K_2, K_5, K_6)
Age (K_1)
Exam (K_3, K_8)
Career routine and inertia (K_2, K_8)
Material perception (K_8)
Deficiencies in using technology (K_3)

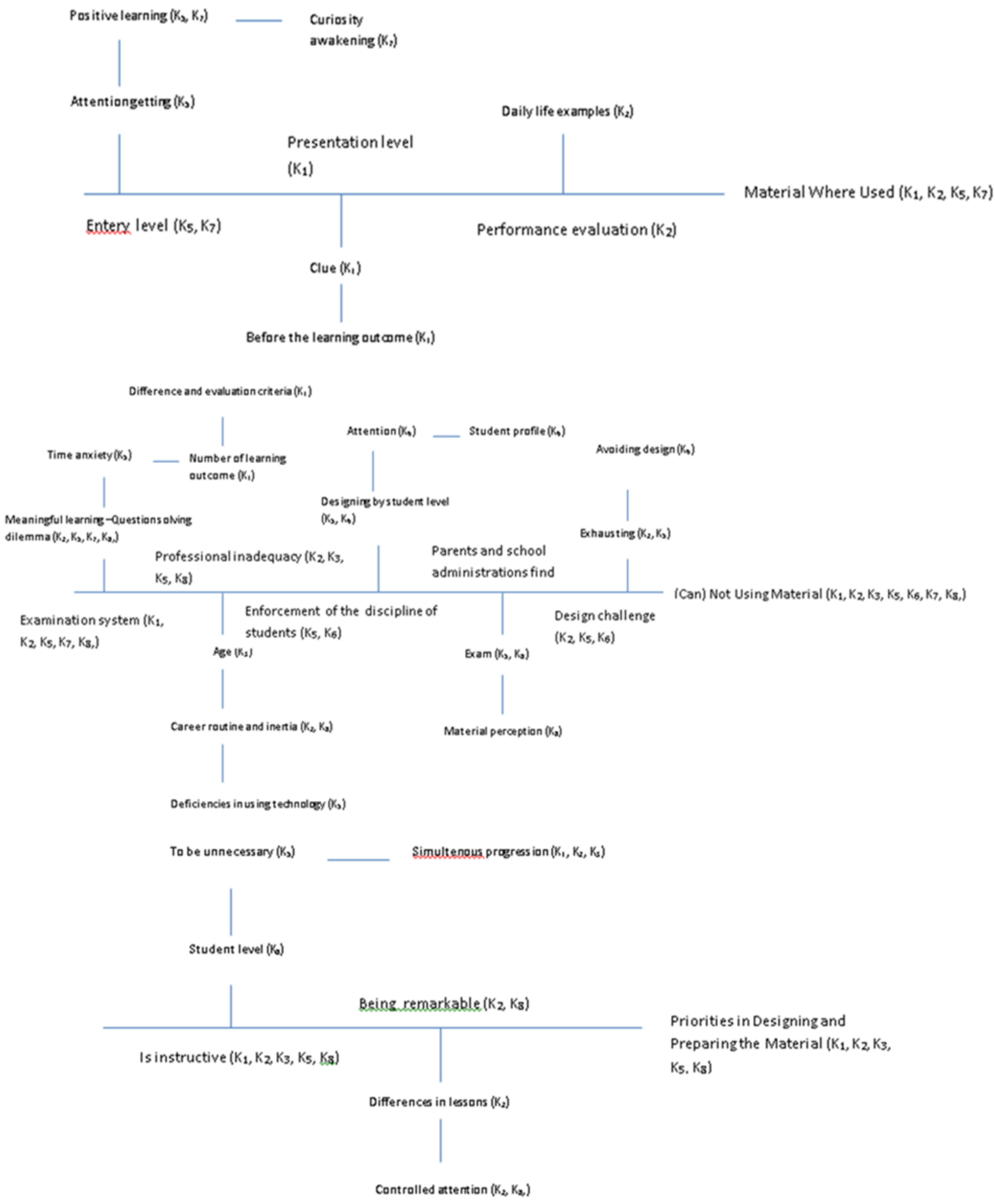


Figure2. Findings regarding the first sub-problem

The interviews revealed that the teachers confused the tools which are benefited from in materials design with the idea of how to put the material in teaching design. K3, one of the participants stated: "*I think material is a compass and a ruler. There are still students who can't use a compass or miter. These should be used in the classroom…*" With these statements, it is clear that the participant cannot understand that teaching how to use teaching tools and the materials that are used in teaching the learning outcomes are different.

K1 mostly emphasized the importance of alternative materials which have simple designs and unique. In this context, K1 stated: "*Actually material is everything a teacher uses. For example, class notes or worksheets… But in general, big and special things are perceived as materials. In fact, everything in the classroom environment is a material. If the aim is to make the students understand, material is anything which is used to make it possible. For example, a piece of paper… I mean, like this… (Here s/he shows that the sum of the interior angles of a triangle is 180° with a triangle shaped paper). I don't know, anything that can clarify the subject. But I couldn't make it clear now… (laughs)*." On the other hand, an alternative material example could not be found from K1's pre-training teaching designs.

Preparation of teaching materials on a theoretical basis enables them to serve the purpose (Burak, 2018). If multiple environment appeals to multiple senses and contribute to the double coding of the content, it may help the individual to learn (Najjar, 1996). Double coding theory of Paivio suggests that stimulants are represented and coded verbally and visually in memory (Paivio, 1986). According to double coding theory, coding a notion in mind both verbally and visually increases the possibility to describe that notion visually and make verbal explanation of it when you encounter it again. Two of the participants stated that using materials in teaching design could be reconstructed even when students forget the notion. K1 explained this situation as: *"…When using materials, even if the student forgets the information, s/he can remember the material and find it herself/himself. I mean it is permanent. In this way, they can don't memorize…*"

A school administration or parents who view using materials in teaching designs as meaningless and a waste of time will demotivate the teacher in designing and using materials and cause the lessons to consist of only doing exercises. Regarding this issue, K8 stated: "*As students prepare for the exams, mathematics teachers implicitly have a mission of solving questions. When you talk about materials, others may think it is nonsense. Firstly the parents complain…*"

K7 stated that using relevant videos as introduction helped create a more sincere athmosphere and continued: *"…I usually make the students watch an interesting video about the topic. Also they get very happy. We get a more positive approach. Then I make them curious. Then I slowly move on to the lesson*." In general, participants took the responsibility to teach and did not create a teaching situation which would allow student to use the material themselves and build the information.

Although the participants had opinions of designing and preparing materials, they could not explain any materials they designed in their teaching designs in the light of these comments during these interviews.

The data from the written documents shows that the materials which the participants thought of using in their pre-training teaching designs were Manipulatives and Visuals. Th teachers were able to design only ten materials for all learning goals.Manipulatives consisted of sub-titles of Physical Manipulatives and Virtual Manipulatives. Additionally, the lack of a B plan in pre-training designs was evident.

Participants' materials using manipulatives in training sessions did not include a detailed planning regarding the aim, or stage of the material. As an example, teaching design of K3 is given below.

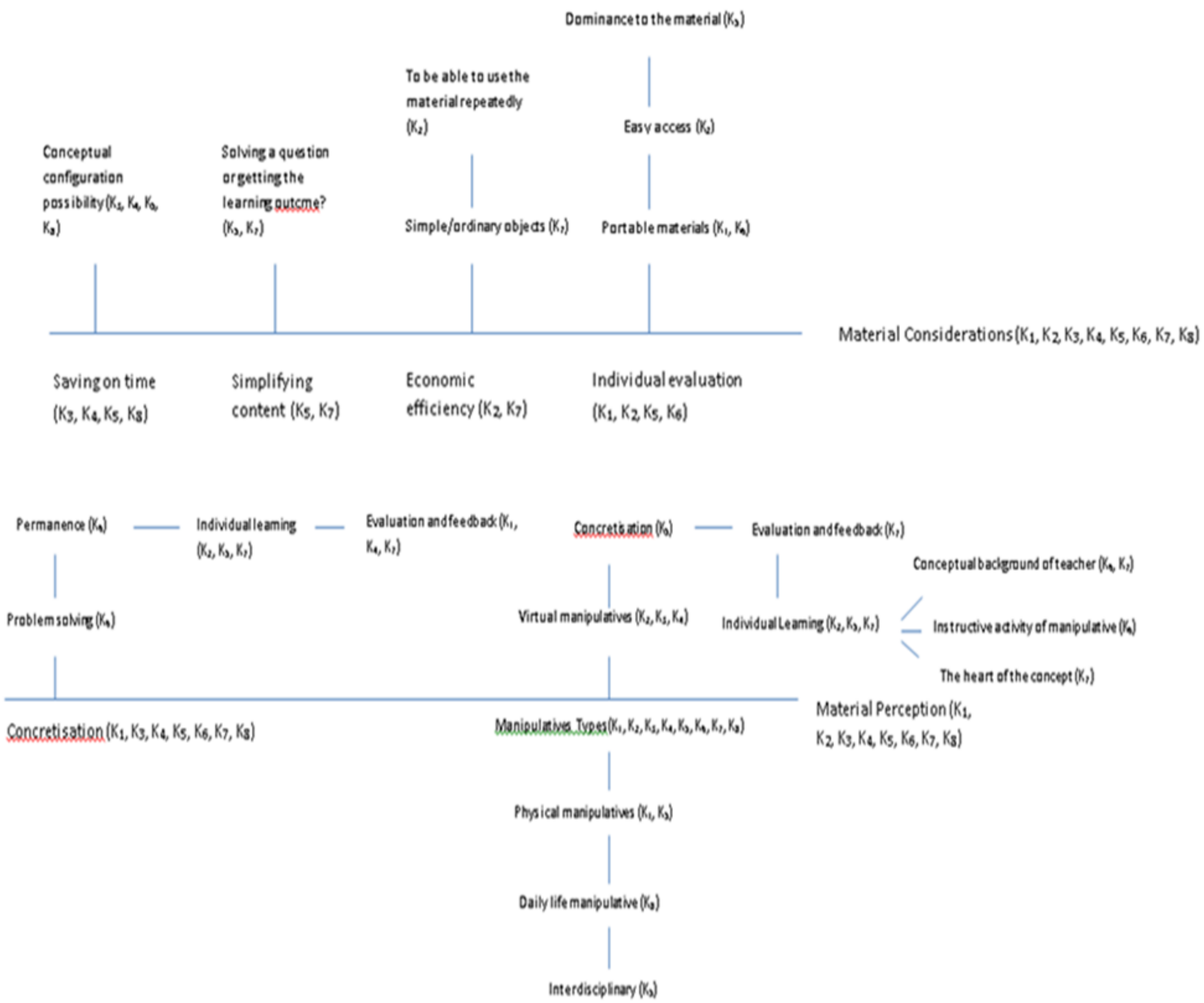
"*I enter the classroom. I show the students the floor of the classroom and ask them how it can be calculated. I teach the concept of circumference on the board. I ask them to calculate the approximate value of the floor. I pose the question "What is area?*" *After the students' feedbacks, I explain the concept. I explain the real life examples of the geometric objects I brought with me. I explain and calculate the area, volume and circumference of each material. Then I explain the difference between actual length and scale length.*"

Additionally, there was not a plan B in the case that students cannot achieve the learning outcomes or learn.

When the teaching designs were examined, it was clear that participants refrained from using computers and other educational technology and most of them were incompetent in using these technologies. K1, who learned how to use GeoGebra in higher education, could not present a construction protocol where the steps were clearly graded.

The basic matter in teaching designs should be to allow effective, deep and meaningful learning of students. In this regard, the participants' materials in pre- training designs could not go beyond visualizing or providing example. Besides, in the interviews with the participants, they stated that the teacher-centered materials (particularly virtual materials) benefitted teachers in terms of time, classroom management and helped them to teach more effectively. Teacher-centered materials are the ones which do not allow students to construct their own learning through trial-and-error. They can only reach conclusion via teacher's instructions. Teacher uses the materials and students passively observe it. All of the materials of the teachers were teacher-centered materials. Moreover, participants stated that students using those materials themselves was a waste of time, distraction and caused false or deficient learning.

Regarding the third sub-problem of the study, five categories were defined. These categories were; Material perception, Assessment of material, Plan B-Assessment of teaching design, (in)ability to use materials and personal assessments. Axial coding regarding these categories and sub-categories are given below:


Dominance to the material
To be able to use the material repeatedly
Easy access
Conceptual configuration possibility
Solving a question or getting the learning outcme?
Simple/ordinary objects
Portable materials
Material Considerations (K1, K2, K3, K4, K5, K6, K7, K8)
Saving on time (K3, K4, K5, K8)
Simplifying content (K5, K7)
Economic efficiency (K2, K7)
Individual evaluation (K1, K2, K5, K6)
Permanence
Individual learning
Evaluation and feedback
Concretisation
Evaluation and feedback
Conceptual background of teacher
Problem solving
Virtual manipulatives
Individual Learning
Instructive activity of manipulative
The heart of the concept
Material Perception (K1, K2, K3, K4, K5, K6, K7, K8)
Concretisation (K1, K3, K4, K5, K6, K7, K8)
Manipulatives Types
Physical manipulatives
Daily life manipulative
Interdisciplinary

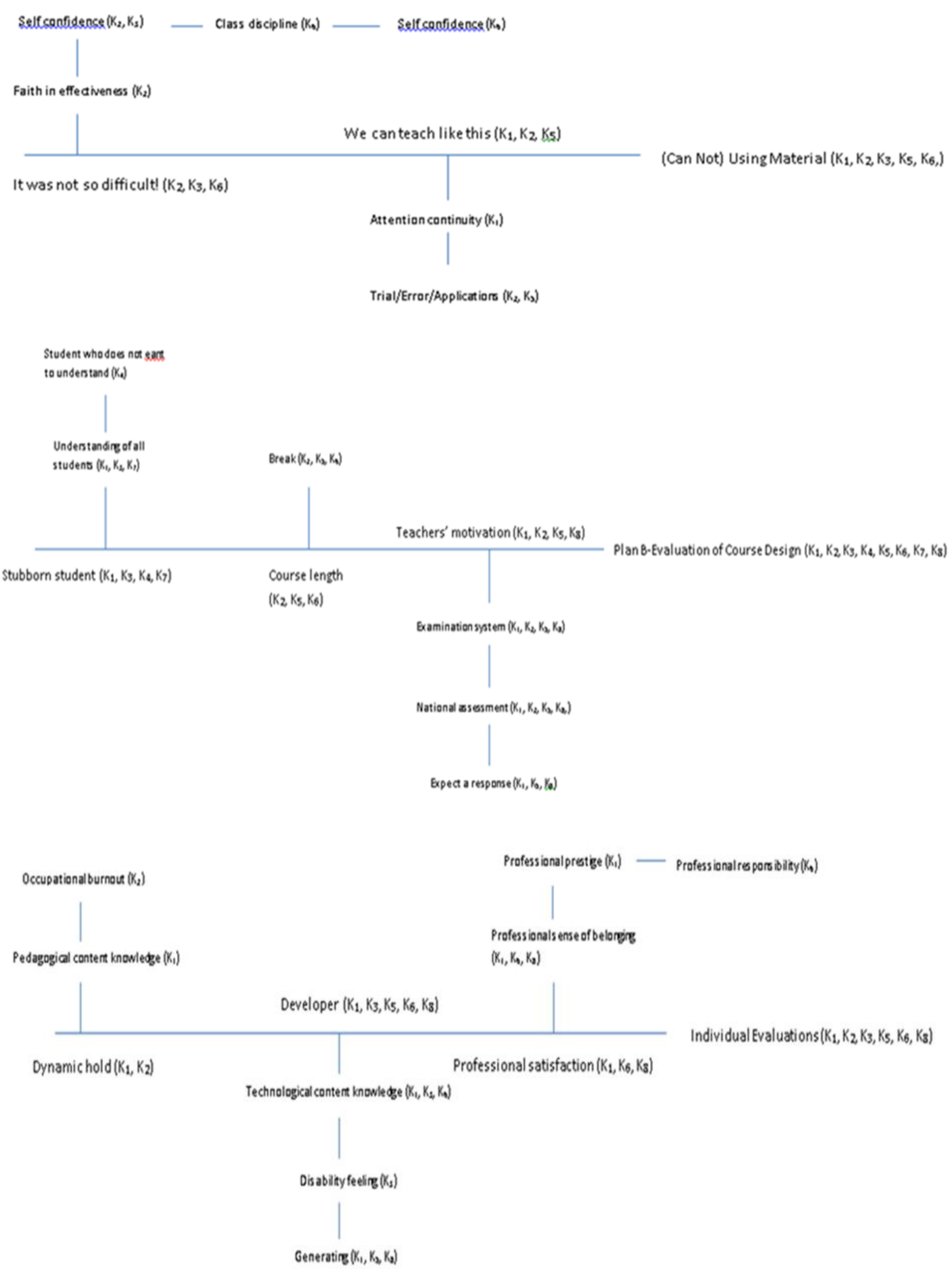


Figure3. Findings regarding the second sub-problem

K7 who emphasized the importance of highlighting the steps of the information that will be learned stated *"…Materials should be designed this way (median material. Actually, I don't* *know how to put it, (ummm) the closer a material is to the heart of a notion, the easier to use it. Because you can see the notion on the material, I mean those steps… Students should not why they take certain steps and construct. In this way, we can prevent them from walking away from the topic…*"

Participants who have examined the materials in teaching designs stated that in these materials, students carried out their learning responsibility by sharing it with the teacher and that this situation differed from the participants' existing material approach. The reflection of awareness is evident in the pre and post training materials.

The participants' statements and teaching designs were examined in general, there are significant differences in their material perceptions; materials categorization, design, preparation and use compared to pre-training. They can now comment on material approach. They can categorize the materials in teaching, assess and criticize the materials they prepare.

K5 who stated there should be emphasis on usefulness in materials design and preparation continued *"…It can be difficult to use these kinds of materials (modular arithmetic board), because your hand can slip or something. But you showed it on median in the second or third lesson. It was good. Then, the sticks were good. I mean, good as in it was clearly set. I mean, it won't go wrong in that minute. Okay, mode was also good. You used it well but I tried. I couldn't manage it as well as you did, for example. I think children can also fail if they tried…*" K5 stated in the detailed interview "*I understood what designing materials is with this training. I mean, it's not like I didn't know anything before that, but…(laughs) What they can use, to what extent? I don't know, where they would have problems? I mean, where is it in the learning outcome, I will highlight it. I didn't know these before. Probably others didn't know either. I am overcome by the urge to teach (laughs)…*" and emphasized the importance of materials in teaching design.

Considering the focus group interviews of the participants and their comments in detailed interviews, they commented constructively, within the scope of material approach, objectively genuinely on the materials used by the researchers in training. This is a considerable change of people who previously had limited assessments regarding materials.

Taking into account that materials facilitates students' learning process, it will be more likely for all or most of the students to learn if they involve in various materials. Whereas, K4 who states that there is a student group who refuses to learn, continues "*It doesn't matter how you teach to a student who doesn't want to understand, s/he will not understand. I think Plan B is unnecessary. You can teach from another angle on the same thing.*"

K6, who thought that materials using in learning-teaching environments will make classroom management more difficult states her/his awareness by saying "*I think all of us love our students very much and that we are very good at mathematics. I mean according to our students… But only loving and being good at mathematics is not enough, unfortunately. How we transfer that information is also very important. You made a presentation, at first. I mean, life environment, receiver language etc… Like system approach. I probably learned it at my undergraduate program but I don't remember it. Why? Because I didn't get any deep understanding on this topic. I mean, it is about how important it is… The deeper and more various the experiences, the more meaningful the learning… But how were we taught about this? Give example from around you and enrich it with concrete objects. But how? I started teaching. I tried to do it the first time. The result was a disappointment. Why? There was not classroom discipline. I don't* *even know why I did what in the design. I mean, there are no steps, to put it your way… (all the participants supports with agreeing mimics)But on a worksheet like this (similarity exercise) with steps suitable for concrete material, why would the student get distracted? Children also want to understand. But as they don't understand, they start to develop prejudice. We also develop them. We suffocate them with our love and knowledge. Our love and knowledge that should be useful cannot go beyond 'Look, also learn this type of questions' unfortunately…*"

K1 who stated that using teaching designs effectively in education environments will change prejudice of both the teachers against their own job and students against their teachers said that training influenced her/him positively with these remarks, *"…I felt special about myself and my job in this process. Students will also respect more because we will make an unclear lesson clear. I think I will be happier then…*" K6 who questioned existing teachings because of the training s/he received stated that s/he was in a stage of having aims for the job by saying "*I started to like making plans. Now I worry about my own lessons (laughs). I mean, it is a little unfair for the children. I started to question myself, what else can I add? My spouse makes fun of me. S/he says that I am not a crazy mathematician (laughs)…*"

When the written documents are examined in detail, it is clear that the participants were more productive at material amount, dividing the teaching design into steps and including a plan B than their first teaching designs. Moreover, the participants used more and various materials. Besides K1's remarks stating "*I look at the one I designed previously and then I look at the latter (Laughs)… Design was hallow in the first one. I mean, ordinary…But the second one… I mean after you… It applies for others too, it seems (they laugh)… Seriously, I designed with a concern to teach. And with that, you aim to teach in those materials. What can I present at what stage, what can I do if they don't understand, etc. It is a good thing. You don't say I showed it and it just happened…*" show that the participants aimed to present a rich teaching content in materials in post-training.

Designs were divided into two main categories of "Manipulatives and Written Materials." Under these titles are;

(1) Manipulatives

(a) Physical manipulatives

(b) Virtual manipulatives

(2) Written Materials

K2 who emphasized the importance of using concrete materials in teaching environments stated "*When the student try it in her/his hand, discuss with the steps in the worksheet, of course s/he will understand better because s/he will construct modular arithmetic. I designed this material in such a way that you can construct it even if you don't know any rules. If you want to solve a watch problem or sum, do it. Students will produce the practice, let them do mathematics in the lesson…*"

Figure4. Concrete manipulative designed by K2 regarding modular arithmetic

The participants benefitted more from virtual manipulatives in post-training when compared to pre-training. However this was still inadequate, considering the recent support and importance in the Fatih Project. Participants were unwilling to use virtual manipulatives. K7 explains this situations by stating "*We have never used it until now. Smartboards came to the class. Okay… tablets programs, all of them are okay. Maybe some our friends have seen it at their undergraduate program but most of us either don't know at all or forgot… At first, we should be taught about the program. We need to practice. How can we do it all of a sudden by ourselves? We criticized Fatih Project. I mean, if it is applied like this, of course it is acceptable… (I'm not sure) if it is meaningful…*"

Whereas K2, K4 and K6's first teaching designs did not include virtual manipulatives, they chose to use these types of materials in their post-training teaching designs. Besides, they followed a staged pattern regarding the use of the manipulatives in their teaching designs.

Conclusions and Suggestions

Teaching materials are not just tools such as pencil, paper, class notes, compass, ruler, etc; a more complex design which are obtained by putting an educational responsibility in these tools (McNeil and Uttal, 2009). While the importance of this situation is known, the concepts of teaching material and tools were not clear for the participants and they used them interchangeably.

When the participants' pre-training teaching designs were examined, it was seen that they included very few materials. As the cause of this situation the participants claimed that they found materials design very challenging, classroom management issues might raise, technological deficiencies and inadequacy regarding materials design/preparation/usage/assessment.

Results revealed that the participants' opinions on not using / not being able to use materials in pre-training were changed in to using / being able to use materials in post-training. In other words, the participants stated that they would use materials more frequently after training. Additionally, they were more confident in materials design and using digital materials.

The pre-training interviews and findings of post-training teaching design show that some of the participants did not interact with software related to teaching and some others applied at their undergraduate program but they did/could not include it in their teaching, and even if they did, they used it to visualize. During and in post-training, the participants could use the software called GeoGebra which was introduced by the researchers and included the accessible applications in their teaching designs. The participants explained this situation with inadequate usage of the smartboards and tablets in Fatih Project. This situations shows the necessity that it should be based on application in the existing teacher training programs and in-service teachers should be supported with in-service training according to developing and changing technology. The participants stated that material usage in teaching will positively affect learning in pre-training. However they experienced inadequacy and avoidance in the process of designing, preparing and using materials. Regarding this issue, it possibly stems from the fact that teachers' level of knowledge on materials usage in teaching is limited to their undergraduate program. This situation causes the teachers to lose the basis, instead of updating it because of them not being able to integrate theoretical knowledge into teaching environments.

Materials of the participants in pre-training were familiar to them from their undergraduate programs. So, it is possible that teacher candidates were not exposed to as many concrete materials and virtual manipulative as possible in the teacher training program.

Teacher candidates need to practice how teaching materials in teaching designs are used. During the study, it was seen that the participants could not arrange how to include concrete materials and virtual manipulatives in teaching designs. In this regard, it is possible that Teaching Technology, Materials Design and Methods of Teaching courses at teacher training programs had deficiencies in teaching and experiencing how to integrate materials into teaching.

The participants included concrete materials more than virtual manipulatives in both pre and post training. The participants stated that they felt inadequate in using virtual manipulatives in the interviews. This situation shows that teacher training programs and in-service training are insufficient at virtual manipulative usage and its integration into teaching design.

The participants who included materials in pre-training included only a teacher-centered, better presentation of the topic instead of allowing students to think and apply. In post-training teaching designs, they were able to integrate the material into teaching design and allow students to try. In this regard, teacher training programs may have deficiencies in providing the teacher candidates with a purpose to use the material.

Ortaöğretim Matematik Öğretmenlerinin Öğretim Materyalleri Hakkindaki Anlayişlari Ve Ürettikleri Materyaller

Özet:

Araştırmanın amacı; matematik öğretmenlerinin öğretim tasarımı sürecinde materyal kullanımlarının incelenmesidir. Eylem araştırmasının nitel paradigmasıyla gerçekleştirilen araştırma, Milli Eğitim Bakanlığı bünyesinde görev yapmakta olan sekiz öğretmenin katıldığı bir eğitim sürecinde gerçekleştirilmiştir. Katılımcılar ile birlikte toplam onbir oturum gerçekleştirilmiş olup veriler yarı-yapılandırılmış görüşmeler, yazılı dokümanlar, video kayıtları ve odak grup görüşmeleri ile toplanmıştır. Veriler içerik analizi ve sürekli karşılaştırmalı analiz yoluyla analiz edilmiştir. Araştırmanın sonucunda; katılımcıların eğitim öncesi öğretim tasarımlarında yer verdikleri materyallerin öğretmen kullanımına açık, öğrencinin yalnızca izleyebildiği yapıda olduğu görülürken; eğitim sonrasında öğrencinin yaparak-yaşayarak deneyimleyebileceği, öğretmenin sadece rehber olduğu materyallere yer verdikleri görülmektedir.

Anahtar kelimeler: öğretmen eğitimi, öğretim tasarımı, manipülatif, sanal materyal.

References

Abrami, P. 2001. Improving judgements about teaching effectiveness using teacher rating forms. *New Directions for Instutional Research, 109*(1), 59-87. https://www.researchgate.net/publication/227630023_Improving_Judgments_About_Teaching_Effectiveness_Using_Teacher_Rating_Forms

Bellinio, J. (2012). *Multi-Sensory Manipulatives in Mathematics: Linking the Abstract to the Concrete*. Yale-New Haven Teachers Institute. http://www.yale.edu/ynhti/curriculum/units/2001/6/01.06.12.x.html (15.02.2016)

Boggan, M., Harper, S, & Whitmire, A. (2010). Using manipulatives to teach elementary mathematics. *Journal of Instructional Pedagogies, 3*(1), 1-6. https://eric.ed.gov/?id=EJ1096945

Bozkurt, A., &Akalın, S. (2010). Matematik öğretiminde materyal geliştirmenin ve kullanımının yeri, önemi ve bu konuda öğretmenin rolü [The place and importance of material development and usage in mathematics teaching and the role of teacher in this subject]. *Dumlupınar University The Journal of Education Faculty, 27*(1), 556-558. https://dergipark.org.tr/download/article-file/55618

Cope, L. (2015). Math manipulatives: Making the abstract tangible. *Delta Journal of Education, 5*(1), 10-19. https://www.semanticscholar.org/paper/Math-Manipulatives%3A-Making-the-Abstract-Tangible-Cope/7ea0bc8c432cea6a5b0411c3148da1e9a329bfdd

Cummings, T., & Worley, C. (1997). *Organization development and change*. Cincinnati: South-Western College.

Dennis, R. (2011). *The effects of the use of manipulatives on the comprehension of math concepts among fifth-grade students*. http://search.proquest.com/docview/915645190?accountid=14771

Domino, J. (2010). *The effects of physical manipulatives on achievement mathematics in grades K-6: a meta-analysis*. Doctoral Thesis, State University of New York: Departmente of Learning anda Instruction Faculty of the Graduate School of the University Buffalo, New York. https://eric.ed.gov/?id=ED522615

Haara, F. (2010). *Unveiling teachers' reasons for choosing practical activities in mathematics teaching*. Doctoral Thesis, University of Bergen, Norway. https://pdfs.semanticscholar.org/5495/969ba63cc03cf5b5f22f0b8759305f6b1547.pdf

Hartshorn, R., & Boren, S. (1990). *Experiential learning of mathematics: Using manipulatives*. ERIC Publications, ERIC Digests. https://eric.ed.gov/?id=ED321967

Huentick, L., & Munshin, S. (2004). *Teaching mathematics for the 21st century methods and activities for grades 6-12*. USA: Pearson and Merril Prentice.

Holmes, A. (2013). Effects of manipulative use on PK-12 mathematics achievement: A meta-analysis. Poster presented at the meeting of *Society for Research in Educational Effectiveness*, Washington, DC.

İnan, C. (2013). Influence of the Constructivist Learning Approach on Students' Levels of Learning Trigonometry and on Their Attitudes towards Mathematics. *Hacettepe University Journal of Eduction Faculty, 28* (3), 219-234. http://www.efdergi.hacettepe.edu.tr/yonetim/icerik/makaleler/219-published.pdf

Kamii, C, Lewis, B., & Kirkland, L. (2001). Manipulatives: When are they useful?. *Journal of Mathematical Behaviour, 20*, 21-31. https://www.sciencedirect.com/science/article/pii/S0732312301000591

Kelly, Ce. (2006). Using manipulatives in mathematical problem solving: A performance-based analysis. *The Mathematics Enthusiast, 13*(2), 184. https://scholarworks.umt.edu/tme/vol3/iss2/6/

Laski, E., Jor'dan, J., Daoust, C., & Murray, A. (2015). *What makes mathematics manipulatives effective? Lessons from cognitive science and montessori education*. SAGE Open, 1-8. http://dx.doi.org/10.1177/2158244015589588

Marshall, L., & Swan, P. (2005). *Developing mathematical thinking with the assistance of manipulatives*. The Mathematics Education into the 21st Century Project. http://www.mathnet.ru/php/conference.phtml?confid=43&option_lang=eng

McClung, L. (1998). *A study on the use of manipulatives and their effect on student achievement in a high school algebra class*. Salem-Teikyo University. https://eric.ed.gov/?id=ED425077

McIntosh, G. (2012). *Testing instrumentation validity for measuring teachers' attitudes toward manipulative use in the elementary classroom. University of Pittsburgh*. http://search.proquest.com/docview/1312423428?accountid=14771

McNeil, N., & Uttal, D. (2009). Rethinking the Use of Concrete Materials in Learning: Perspectives From Development and Education. *Child Development Perspectives, 3*(3), 137-139. https://onlinelibrary.wiley.com/doi/full/10.1111/j.1750-8606.2009.00093.x

The Ministry of National Education. (2008). *Öğretmen yeterlilikleri [Teacher qualifications]*. Ankara: Devlet Kitapları Müdürlüğü.

Moyer, P. (2001). Are we having fun yet? How teachers use manipülatives to teach mathematics. *Educational Studies in Mathematics, 47*(1), 175-197. https://digitalcommons.usu.edu/cgi/viewcontent.cgi?article=1054&context=teal_facpub

Munger, D. (2007). *Science blogs. Children learn and retain math better using manipülatives*. scienceblogs.com/cognitivedaily/2007/10/09/children-learn-and-retain-math/

Moyer, P., Bolyard, J., & Spikell, M. (2002). What are virtual manipulatives?. *Teaching Children Mathematics, 8*(6), 372-377. https://www.researchgate.net/publication/43952500_What_are_Virtual_Manipulatives

Ojose, B., & Sexton, L. (2009). The effect of manipulative materials on mathematics achievement of first grade students. *The Mathematics Educator, 12*(1), 3-14. https://www.researchgate.net/publication/265939371_The_Effect_of_Manipulative_Materials_on_Mathematics_Achievement_of_First_Grade_Students

Patton, M. (1987). *How to use qualitative methods in evaluation*. Newbury Park, CA: Sage.

Perry, B., & Howard, P. (1994). *Manipulatives: constraints on ınstruction*. https://www2.merga.net.au/documents/RP_Perry_Howard_1994.pdf

Shaw, J. (2016, 11 17). *Manipulatives enhace the learning of mathematics*. Education Place. http://www.eduplace.com/state/author/shaw.pdf.

Thompson, P. (1992). Notations, conventions, and constraints: Contributions to effective uses of concrete materials in elementary mathematics. *Journal for Research in Mathematics Education*, *23*(2), 123-147. http://dx.doi.org/10.2307/749497

Thompson, P.. (1994). Concrete materials and teaching for mathematical understanding. *Arithmetic Teacher, 41*(9), 556-558. https://www.researchgate.net/publication/264119418_Concrete_materials_and_teaching_for_mathematical_understanding

White, K. (2012). *The effect of an instructional model utilizing hands-on learning and manipulatives on math achievement of middle school students in Georgia*. Lynchburg, Virginia: Liberty University, Lynchburg, VA.

Yalın, H. (2015). *Öğretim teknolojileri ve materyal geliştirme [Instructional technologies and material development]*. Ankara: Nobel.

Yıldırım, A., & Şimşek, H. (2006). *Sosyal bilimlerde nitel araştırma yöntemleri [Qualitative research methods in the social sciences]*. Ankara: Seçkin.

Appendix

Appendix A. Semi-structured Interview Form to Identify the Problem Case

Hello,

As part of my research, we are planning to conduct a study on our teachers' approach, usage, ability to design, prepare, develop and present materials. At this point, assessments which we will carry out with you are quite important to reveal the problem case clearly. The interview will be solely used for scientific purposes and recorded with assurance that it will not be shared anywhere else. If you consent to it, we can begin.

(1) Can you describe materials?

(2) What is the place of materials in mathematics education?

- If there should be materials, explain why.
- If there should not be materials, explain why.

(3) Do you use materials in your teaching?

- If you use them; for what purpose, what type of materials and how often do you use them?
- If you do not; why not?

(4) Can you make an evaluation of materials usage in our country's mathematics education?

Thank you.

Appendix B. Interview Form

Hello,

Dear teachers, as part of my research, your opinions on materials and materials in mathematics teaching are of great value to me. Therefore, this interview will be solely used for scientific purposes and recorded with assurance that it will not be shared anywhere else. If you consent to this, we can begin our interview.

(1) What does materials usage mean to you?

(2) What are your opinions on materials usage in educational environment?

(3) What re your views on materials usage in mathematics teaching?

(4) Do you use materials in your lesson designs? If you do/ do not, why/ why not?

(5) What is the number of materials you used in a teaching design?

(6) What type of materials do you use?

(7) Is there a material that is unique to you?

(8) What are the steps you follow in designing a material?

(9) What are the things you pay attention to while designing a material?

(10) When do you use the material in your teaching design?

(11) What do you do when you cannot achieve your goal with the material?

Appendix C: Focus Group Interview Form

(1) How do you design your teaching for this learning outcome? Explain.

(2) What do you think about this design? Share.

(3) What are your views on the materials in the design? Explain.

(4) How do you assess this material from a student's point of view? Explain.

(5) How do you assess this design from a teacher's point of view? Explain.

Necatibey Eğitim Fakültesi Elektronik Fen ve Matematik Eğitimi Dergisi
Cilt 16, Sayı 2, Aralık 2022, sayfa 326-353. ISSN: 1307-6086
Necatibey Faculty of Education Electronic Journal of Science and Mathematics Education
Vol. 16, Issue 2, December 2022, pp. 326-353. ISSN: 1307-6086

Research Article

The investigation of middle school students' entrepreneurial skills in terms of Entrepreneurship-Based STEM Education: A mixed method study

Meryem Meral [1], Sema Altun Yalçın [2]

[1] Graduate School of Natural and Applied Sciences, Erzincan Binali Yıldırım University, Erzincan, Turkey, meryemmeral96@gmail.com, https://orcid.org/ 0000-0001-6256-1725

[2] Graduate School of Natural and Applied Sciences, Erzincan Binali Yıldırım University, Erzincan, Turkey, saltun_11@hotmail.com, https://orcid.org/0000-0001-6349-2231

Received : 22.06.2022 Accepted : 07.12.2022

Doi: 10.17522/balikesirnef.1133770

Abstract –

This study was carried out to investigate middle school students' entrepreneurial skills in terms of E-STEM education. The universe of the study comprised of middle school students in a rural area. The sample of the study, in which single group pre-and post-test design was used, consisted of 20 eighth-grade students. The mixed method was preferred. The entrepreneurship scale was used to obtain quantitative data. Semi-structured interview form was used to obtain qualitative data. E-STEM education was given to the students for 8 weeks. Pre-and post-test means of quantitative data were compared by paired-sample t test and content analysis method was used to analyze qualitative data. As a result, the scores of students' entrepreneurial skills and sub-dimensions of entrepreneurship increased in favor of the post-test, although the overall increase was not significant. In addition, according to qualitative data, it was observed that E-STEM activities had a positive effect on the sub-dimensions.

Key words: entrepreneurship, middle school students, STEM activities

----

Corresponding author: Meryem MERAL, meryemmeral96@gmail.com

Acknowledgement: This study has been produced from the author's master's dissertation. The effect of STEM activities performed via simple materials on students' entrepreneurship and self-regulation skills by Meral, Meryem, Master Dissertation, Erzincan Binali Yıldırım University, 2020.

Introduction

Education has gained a new definition in terms of its purpose and function in 21st-century society. The reason for this change stems from the fact that the developments in today's society and economy require people who are equipped and qualified with certain skills (Ananiadou & Claro, 2009). Skills that individuals require in our age are included in "21st- century skills" (Johnson, 2009). "Entrepreneurship" is one of these 21st-century skills included in the category of "career and life skills" (Trilling & Fadel, 2009). Hindle and Rushworth (2000) have asserted that entrepreneurship is an activity that is a fundamental factor in the creation and management of a new organization designed in order to create a unique and innovative opportunity. It has been mentioned that statements defining entrepreneurship focus on using the skills that characterize entrepreneurial individuals and the processes that play a part in entrepreneurship (Nafukho et al., 2010). Rocha and Birkinshaw (2007) have drawn attention to the discourse that an entrepreneur is defined as the owner of a small or medium-sized company or a new businesses. Some studies concerning entrepreneurship have indicated that entrepreneurship has been regarded as a key element of rising living standards (Azoulay et al., 2018). However, it has been argued that successful entrepreneurship is rare among societies and the vast majority of entrepreneurs fail to deliver major or creative innovations (Azoulay et al., 2018). Therefore, more emphasis should be placed on the development of entrepreneurial characteristics of individuals in order to prevent this failure and to create an innovative society.

In studies conducted on this subject, students were given entrepreneurship training for the purpose of promoting entrepreneurial skills and increasing the number of entrepreneurial individuals and it has been observed that these trainings have contributed to the entrepreneurial skills of the students. Lepuschitz et al. (2018) conducted a pilot study in two schools in Vienna, Austria for the purpose of the development of entrepreneurial skills in middle schools. Some activities including robotics, marketing, business management, problem solving, teamwork were carried out by the students within the scope of this entrepreneurship education. It has been speculated that these practices could improve the entrepreneurial skills of middle school students. Another research conducted by Sanchez (2013) has focused on the effect of entrepreneurship education on students' desire to start a new job. This entrepreneurship program consists of four basic components:(1) basic teachings of management and marketing adapted to the age of the students; (2) practices in skills such as self-efficacy and risk taking; (3) business plan and (4) interaction with practice and, consequently, a significant increase has been observed in favor of the experimental group and revealed that entrepreneurship education positively affected students' desire to start a new job. Moreover, in another study conducted by Uygun and Güner (2016), university students were informed about the characteristics of entrepreneurial individuals, entrepreneurship culture, innovation and creativity, as well as details of business establishment and financial planning during the 14-week training period. The results of this study have indicated that entrepreneurship education contributes to the formation of positive results in terms of entrepreneurial knowledge and skills. In addition, Avcı (2018) has examined the effect of differentiated teaching practices, which include skills such as risk taking, teamwork, creativity, critical thinking, and problem solving, on students' entrepreneurial skills and a significant increase in entrepreneurial skills of the experimental group has been found. In addition, the acquisition of entrepreneurial skills at early ages is considered to be a serious issue besides many other skills (Tarhan & Kılıç, 2017). Therefore, it was considered that the idea of integrating the education given to primary school students with entrepreneurial skills would be appropriate. STEM education which aims to develop 21st-century skills and entrepreneurial competencies was preferred for the implementation of the research in accordance with the purpose. The term "STEM education" refers to the teaching and learning process in the fields of Science, Technology, Engineering and Mathematics (Gonzalez & Kuenzi, 2012). STEM education also has common goals with the 21st century society in terms of being an instructional model that enables individuals to develop entrepreneurial and collaborative work as well as high-level skills (Walan, 2019). In addition, Deveci (2018) stated that STEM education and the concept of entrepreneurship which is one of the 21st century skills have been emphasized recently. Furthermore, the inclusion of the concept of entrepreneurship under the name of "life skills" in the new middle school science curriculum has indicated that the goals of the new curriculum have a common ground with the STEM approach (Deveci, 2016). STEM education has a sub-field called E-STEM (Entrepreneurship, Science, Technology, Engineering, Maths) (Caldwell et al., 2018). It is aimed to acquire entrepreneurial skills as well as other acquisitions via this field which is composed of the integration of entrepreneurship and STEM education (Ezeudu et al., 2013). In addition to the main goals of STEM education, orienting individuals to professions that include entrepreneurship has been one of the goals of this sub-field (Langdon et al., 2011). One of the main reasons for this situation has been the increase in the rate of occupations involving STEM fields in recent years (Caldwell et al., 2018). According to Farwati et al. (2021), teachers implement STEM education to develop students' entrepreneurial skills and various

21st century skills. Ahmad and Siew (2021) emphasized the importance of integrating entrepreneurial thinking into STEM education. Shahin et al.(2020) investigated the effects of STEM-based entrepreneurship program on entrepreneurial intention of female students. The findings of the study indicated that the entrepreneurship program developed female students' entrepreneurial attitudes, skills such as problem solving, creativity, leadership, and risk taking. Thus, it is essential for individuals who will choose E-STEM-related professions in the future to gain knowledge and experience in E-STEM fields so that individuals could have the opportunity to get a job and start a business. In this context, it is possible for students to develop both their entrepreneurial characteristics and engineering skills via E-STEM education (Deveci et al., 2015).

The research is significant in terms of focusing on STEM education and entrepreneurial skills which have been the issues coming to the fore in recent years. Moreover, despite the fact that there are many STEM related studies focusing on problem solving, critical and creative thinking skills (Çakır, 2018; Öztürk, 2018; Topsakal, 2018; Aydın, 2019; Özkızılcık & Cebesoy, 2020), the number of studies examining the effect of STEM education on entrepreneurial skills is relatively limited (Konuş, 2019; Deveci, 2018). In addition, the acquisition of entrepreneurial skills is highly vital as in today's society due to the fact that individuals are expected to establish their own businesses, produce creative solutions to problems, act with confidence and be aware of themselves. Therefore, individuals need to be taught and trained in this field in order to acquire entrepreneurial skills. Because of the fact that these features and the possibility of emergence of these features are closely related to entrepreneurial skills of individuals. In addition, Nicolaides (2011) drew attention to the fact that entrepreneurship is an important factor especially for the developing societies. Stevenson (2000) stated that enterprises play a key role in establishing a business and the importance of this skill has increased in the rapidly developing technology age. Aytaç (2006) suggested that entrepreneurial skills are an advantageous feature for children in terms of self-sufficiency and self-confidence. Besides, Abbasi et al. (2011) argued that entrepreneurial skills play a key role in individuals in terms of affecting leadership and communication skills as well. In this context, it can be inferred from these statements that entrepreneurial skills are also necessary for the socialization of students. Development of entrepreneurial skills enable students to create a certain profile in their future lives. Daniel et al. (2017) drew attention to the impact of entrepreneurship on economic independence and claimed that this impact may be able to help individuals establish their own businesses or find a good job in the future, provided that the education on entrepreneurial skills can be effective. Based on these assertions, it is suggested that more attention must be paid to the studies on entrepreneurial skills and the factors affecting students' entrepreneurship. Thus, this study aimed to examine the entrepreneurial skills of middle school students in terms of entrepreneurship-based STEM education carried out with simple materials. In addition, determining the entrepreneurial perceptions of students before and after the /implementation is one of the aims of the research. Besides, it is also aimed to provide middle school students with a basic perspective on STEM education with cost-effective and easily available materials. Thus, answers to the following questions were sought in the study.

1. Does entrepreneurship-based STEM education have an impact on middle school students' entrepreneurial skills?
2. What are the middle school students' entrepreneurial perceptions before and after the implementation?

Method

Study Sample

The universe of the study comprised of middle school students being educated in a rural area and the study sample consisted of 20 eighth-grade rural students ranged in age from 13 to 14 years. The research was conducted in the fall and spring semesters of the 2019-2020 academic year. Purposive sampling method was used in sample selection. The purposive sampling method is the determination of the study sample based on specific objectives or qualities (SA et al., 2021). The qualities considered in this research are "homogeneity" and "convenience". Purposive sampling method enables the sample selections including groups that are easily accessible (Nartgün & Kaya, 2016) and homogenous in terms of age, culture and socio-economic status (Etikan, 2016).

Research Design

"Mixed method" in which qualitative and quantitative methods are both used was preferred in the study. It is seen that the mixed method is interpreted differently by many researchers. While Venkatesh (2013) and Leech et al. (2010) simply define the mixed method as a method created by the combination of qualitative and quantitative designs, for some researchers, the mixed method is more than a simple combination of two approaches due to its integrative role in study results (Morse & Cheek, 2015). Caruth (2013) suggested that mixed method can offer broader insights and generate more knowledge. These features are mainly related to the pragmatic nature of the mixed method (Hall, 2013). Therefore, the mixed method was chosen in this study in order to better explain the quantitative data and to enhance the perspective and understanding of the study results. The mixed method basically includes three different designs that are named as; convergent design, sequential-exploratory design and sequential-explanatory design (Creswell, 2013). Explanatory design was used in this research, since quantitative methods were applied first, and then qualitative methods were used to confirm the quantitative data and to explain them in more detail (Ivankova et al., 2006). Single-group pre-and post-test model, which is one of the weak experimental designs, was used for the implementation of the research. Experimental designs are based on holding all conditions and variables constant except the independent variable in the process of the research and examining the effect it has on the experimental group (Ross & Morrison, 2004). A single group pre-test-post-test model was used to obtain the quantitative data of the study. This model is used to compare groups and to determine the change resulting from the experimental intervention (Büyüköztürk et al, 2016). Interview technique was used to obtain qualitative data. Interviewing is an indispensable and important data collection technique in social sciences research (Briggs, 1986). People are provided with information about their own feelings, attitudes and experiences through interviews (Türnüklü, 2000).

Data collection

"Entrepreneurship Scale for Middle School Students" developed by Özcan (2019) was used in order to collect the quantitative data of the study. This scale was a five-point Likert-type one including strongly agree, agree, don't know, disagree and strongly disagree choices and applied to students before and after the experimental application. A semi-structured interview consisting of seven questions in parallel with the items in the sub-dimensions of the entrepreneurship scale was also applied before and after the application for the purpose of obtaining the qualitative data and the interview questions were asked to determine the entrepreneurial skill levels of the students. The interview questions which were asked to determine entrepreneurial skills and their relevant sub-dimensions (self-confidence, the perception of innovation and creativity, leadership and tendency to stand out, social skills and group work, risk taking tendency) are as follows:

(1) "Can you freely share your opinion you think is correct with your friends?" "self-confidence".

(2) "What is it like for you to stand out during group work?" "Leadership and tendency to stand out".

(3) "What do you think about group leadership?" "Leadership and tendency to stand out".

(4) “Can you find creative solutions to the problems you face?” “the perception of innovation and creativity”.

(5) “Do you feel like you are doing wrong while determining your ideas or decisions?” “Self-confidence”.

(6) “What do you think about group work with friends?” “Social skills and group work”.

(7) “Can you take risks in situations that you think are important?” “Risk taking tendency”.

Data Analysis

“Paired samples t test” which is one of the statistical methods was used for the analysis of quantitative data. Paired samples t-test is an analysis method used in the application of a test-retest situation and in the process for investigating the relationship or the level of difference between pre-test and post-test scores (Mee & Chua, 1991). Qualitative data were analyzed by using the “content analysis method”. Content analysis provides the researcher with the opportunity to analyze data with an impressionist, instinctive and interpretative approach (Hsieh & Shannon, 2005). Applications in the studies conducted by Maxwell (1992), Lincoln and Guba (2000), Creswell and Miller (2000) were used in order to validate the qualitative measurement tools and the analysis of the research. According to these studies, in order to ensure the validity of the qualitative measurement tools, it was ensured that the study group gave sincere and correct answers to the questions. The interviews were recorded and transformed into written texts in order that the researcher could objectively quote the answers given to the questions. In this context, measurement tools were examined by professionals in the field of educational sciences, and the questions in the measurement tool were finalized by taking into consideration the opinions of the experts.

Validity and Reliability

Şencan (2005) have stated that validity is a concept used to determine to what extent the questions in the measuring tool represent the characteristics aimed to be measured. In this respect, the qualitative data of this research represent the characteristics that are aimed to be measured as they are prepared in accordance with the sub-dimensions of the entrepreneurship scale. Moreover, in order to ensure the reliability of the qualitative data of the research, codes and categories determined independently by two different researchers were created and compared within the scope of the content analysis. The number of consensus and disagreement was determined and the reliability of the study was ensured. The reliability coefficient was determined according to the formula (Reliability =(Agreement / Agreement + Disagreement x100) developed by Miles and Huberman (1994). The reliability coefficient determined in this study was found as 91% [(42 / 42+5)x100] for the interview questions prepared for entrepreneurial skills.

Implementation

STEM education was given to the students in the experimental group for 8 weeks by the applications involving entrepreneurship-based STEM activities each week within the scope of the application process. Research data were collected by using qualitative and quantitative measurement tools. The studies on the entrepreneurship-based STEM education were examined and the STEM activities were performed by considering the developmental levels and readiness of the experimental group. The “E-STEM” model which was discussed and explained by Deveci in the book called “STEM Education from Theory to Practice” was taken as basis in the application steps of the activities that are economical in terms of time and cost. In this regard, the acquisitions for “the catapult” activity were determined under the name of “Science, Engineering and Entrepreneurship Practices” and the necessary theoretical information about this activity was provided to the students. Then, they were expected to define the problem by associating it with a need in daily life. A discussion and brainstorming environment was created, for instance; “Let’s assume there is a problem such as throwing an object away. So how would we deal with this problem?”. Afterward, students were given the chance to answer the question “What kind of design can we develop to solve this problem?” and subsequently, they were asked to choose the simple materials with which they could model this design. Each group of students discussed how effective the work was in solving the given problem by designing their joint work, and also the questions “What variables must be changed?” (For example, if the rubber bands are wrapped more in the tongue bar, can the object be thrown further? etc.) were asked by the students and they argued on which principles of STEM fields they used to develop this work. Finally, students were asked to compete and market their work for the purpose of putting more emphasis on their entrepreneurial skills.

Findings

The Analysis of Quantitative Data

The quantitative data of the study were analyzed by statistical methods. It was observed that the obtained quantitative data were normally distributed by considering Shapiro-Wilk test due to the fact that the sample number was below 50 (Razali and Wah, 2011). According to the Shapiro-Wilk test, A p-value higher than 0.05 (> 0.05) indicates that the data group is normally distributed (Taşpınar, 2017). The pre-and post-test results of the “Entrepreneurship

Scale for Middle School Students" and its sub-dimensions applied to 20 students before and after the activity were analyzed with the "Paired Sample t-test" method. The findings obtained are shown in Table 1 and Table 2.

Table 1. Results of the Paired Sample t-Test on entrepreneurship scale

Measurements	N	x̄	Sd	T	df	P
Pre-test	20	96.95	19.10215			
				-1.692	19	.107
Post-test	20	107.20	17.67127			

The paired sample t-test results of the pre-and post-test scores of students' entrepreneurial skills are shown in Table 1. Paired samples t-test was conducted to determine whether there was a significant difference between pre-test and post-test scores. Since the significance level was higher than 0.05 ($p > 0.05$), no significant difference was found between pre-and post scores.

Table 2. Results of the Paired Sample t-Test for the sub-dimensions of the entrepreneurship scale

Sub-dimension	Measurements	N	x̄	Sd	T	df	P
Self-confidence	Pre-test	20	16.45	6.39469			
					-.575	19	-.572
	Post-test	20	17.60	4.66115			
The perception of innovation and creativity	Pre-test	20	22.35	5.8063			
					-.434	19	.669
	Post-test	20	23.15	5.7608			
Leadership and tendency to stand out	Pre-test	20	21.5	3.85937			
					-.862	19	.400
	Post-test	20	22.5	2.74341			

Social skills and group work	Pre-test	20	24.15	4,35618			
					-1.663	19	.113
	Post-test	20	26.25	4.15331			
Risk taking tendency	Pre-test	20	15.85	4.53379			
					-1.228	19	.234
	Post-test	20	17.70	4.71392			

The paired sample t-test results of the pre-and post-test scores of the students' entrepreneurial skills are shown in Table 2. Since the significance level is higher than 0.05 ($p > 0.05$) for each sub-dimension, it can be stated that there is no significant difference between pre-post scores.

The Analysis of Qualitative Data

The qualitative data of the study were analyzed by the content analysis method. In this context, codes and categories were created for the answers to the 7 questions that were asked in order to measure the entrepreneurial skills of the students, and the frequency and percentage rates of these codes were shown as pre-and post-test along with the necessary explanations given under the tables.

Table 3. Student views on the first question

Category	Code	Pre-test Frequency (F)	Post-test Frequency (F)	Pre-Test (%)	Post-test (%)
Yes	Freely	1	4	5	20
	Directly	12	12	60	60
	Unsure	1	1	5	5
No	Fear	1	0	5	0
	Hesitation	5	3	25	15
Total		20	20	100	100

In Table 3, the answers to the question "Can you freely share your opinion you think is correct with your friends?" consisted of two categories called yes and no, and 5 five codes (freely, directly, unsure, fear, hesitation). The category of yes consisted of codes that are called freely, direct, unsure. The frequency value of the code called "freely" increased in favor of the post-test, while frequency values did not change in other codes in this category. The codes in this category were composed of the thoughts that students could freely or directly express their ideas even if they were not sure that they were correct. The category of no consisted of the codes that are called fear and hesitation. The frequency values of these codes decreased in favor of the post-test. The codes in this category comprised of the thoughts that the students were afraid of expressing their ideas freely and they avoided expressing their thoughts because of being a person with a timid personality or assuming that they would feel embarrassed about expressing opinions, asking questions about the point they do not understand, and thinking that they would be ridiculed by their friends. Some of the students' thoughts are given below:

S1: "...I can express my ideas, even when I am not sure…"

S2: "...No, I can not. Because, I feel afraid of being mocked if I am wrong…"

Table 4. Student views on the second question

Category	Code	Pre-test Frequency (F)	Post-test Frequency (F)	Pre-Test (%)	Post-test (%)
Liking	Pleasant	10	13	50	65
Disliking	Timidity	8	5	40	25
	Unassociated	2	2	10	10
Total		20	20	100	100

According to Table 4, the answers to the question "What is it like for you to stand out during group work?" consisted of two categories called liking and disliking, and three codes (pleasant, timidity, unassociated). The category of liking consisted of a single code called pleasant. The frequency value of the code called "pleasant" increased in favor of the post-test. This code was composed of the thoughts that students liked to stand out during group work and they did not hesitate to stand out. The category of disliking consisted of the codes that are called timidity and unassociated. The frequency value of "timidity" decreased in favor of the post-test while the frequency level of "unassociated" remained stable. The codes in this category comprised of the thoughts that students could not highlight themselves and prefer to stay in the background. Moreover, students stated that many disagreements could arise during group work, they could more freely make their own decisions, they did not prefer to make decisions on behalf of the group, also they did not want to take the responsibility of their groupmates. Some of the students' thoughts are given below:

S3: "...I don't like standing out as a single person…"

S4: "...I got used to standing out thanks to the activities…"

Table 5. Student views on the third question

Category	Code	Pre-test Frequency (F)	Post-test Frequency (F)	Pre-Test (%)	Post-test (%)
Affective	Liking	6	10	12	20
	Desire	6	10	12	20
	Disliking	14	10	27	20
	Hate	14	10	27	20
Self-efficacy	Ability	1	3	2	5
	Disability	10	8	20	15
Total		51	51	100	100

In Table 5, the answers to the question "What do you think about group leadership?" consisted of two categories called affective and self-efficacy, and six codes (liking, desire, disliking, hate, ability, disability). The category of affective consisted of codes that are called liking, desire, disliking, hate. The frequency value of these codes remained unchanged. The codes in this category were composed of the thoughts that some students loved and played a role in group leadership, while some of them did not like leadership because they could not highlight themselves within their group, and also the students preferred group leadership because they liked to organize. The category of self-efficacy consisted of the codes that are called ability and disability. The frequency values of these codes changed in favor of the post-test. The codes in this category comprised of the thoughts that some students' personality were not agreeable for group leadership because these students could not bring themselves to the fore and group leadership is not a responsibility that they could take because of the fact that they did not trust themselves in controlling the group and also some students belived that they were able to lead the group through the activities. Some of the students' thoughts are given below:

S5: "...I don't prefer to be a group leader, it's not something easy for me…"

S6: "...I don't want to be a group leader because I don't want to take that responsibility…"

Table 6. Student views on the fourth question

Category	Code	Pre-test Frequency (F)	Post-test Frequency (F)	Pre-Test (%)	Post-test (%)
Yes	Comprehension	3	3	12	10
	Test questions	2	2	8	6
	Relationships	5	5	19	16
	Technical problems	8	8	30	26
	Demanding help	5	5	19	16
	STEM activities	0	5	0	16
No	Disability in problem solving	3	3	12	10
Total		26	31	100	100

In Table 6, the answers to the question "Can you find creative solutions to the problems you face?" consisted of two categories called yes and no, and seven codes (comprehension, test questions, relationships, technical problems, demanding help, STEM activities, disability in problem solving). The category of yes consisted of codes that are called comprehension, test questions, relationships, technical problems, demanding help, STEM activities. The frequency values of the codes remained unchanged except for the code "STEM activities". The frequency of the code "STEM activities" increased in favor of post-test. The codes in this category were composed of the thoughts that students could solve problems that arise within the family of friends' circle or in technological devices and also they could cope with problems by receiving help from others. In addition, they could solve problems about doing homework, tests or about mathematical problems. Some students were able to repair broken devices at home, whereas a couple of students used internet search to solve their problems. Apart from the most students, five students stated that they encountered and solved some problems during the implementation of the activities. The category of no consisted of the code "disability in problem solving". The frequency values of this code remained equal. The codes in this category comprised of the thoughts that students could not figure out the right way of problem solving, even though they tried to solve the problem and they could not think of any creative solutions because they only focused on the problem instead of its solution. Some of the students' thoughts are given below:

S7: "...I designed the parachute using balloons during a STEM activity and it worked…"
S8: "...My solution never works whenever I try to solve a problem…"

Table 7. Student views on the fifth question

Category	Code	Pre-test Frequency (F)	Post-test Frequency (F)	Pre-Test (%)	Post-test (%)
Yes	Strongly yes	9	9	45	45
	Sometimes	5	5	25	25
No	Strongly no	6	6	30	30
Total		20	20	100	100

In Table 7, the answers to the question "Do you feel like you are doing wrong while determining your ideas or decisions?" consisted of two categories called yes and no, and three codes (strongly yes, sometimes, strongly no). The category of yes consisted of codes that are called strongly yes and sometimes. The frequency values of the codes remained unchanged. According to the statements of this category, some students were not certain in decision making because of feeling as though their decisions were wrong. These students also stated that they felt like they were wrong while answering a question. The category of no consisted of the code "strongly no". The frequency values of this code also remained the same. The codes in this category comprised of the thoughts that students were mostly satisfied with the decision they make, they did not feel like they were making a wrong decision because they generally trusted in their decision, and they had the right information about their ideas. Some of the students' thoughts are given below:

S9: "...I sometimes feel like I give wrong answers to the questions in the classroom…"
S10: "...I am usually confident with the decisons I make…"

Table 8. Student views on the sixth question

Category	Code	Pre-test Frequency (F)	Post-test Frequency (F)	Pre-Test (%)	Post-test (%)
Affective	Liking	8	15	36	57
	Disliking	8	2	36	8
	Inconstancy	4	3	18	12
	Timidity	1	1	5	4
Self-efficacy	Suitability	1	5	5	19
Total		22	27	100	100

In Table 8, the answers to the question "What do you think about group work with friends?" consisted of two categories called affective and self-efficacy, and five codes (liking, disliking, inconstancy, timidity, suitability). The category of affective consisted of codes that are called liking, disliking, inconstancy, timidity. The frequency values of the codes changed in favor of post-test, while the code of "timidity" remained unchanged. According to the statements of this category, students liked and felt better in group work after the implementation, and also they liked learning more because they had fun with their friends. Moreover, they could get help from their friends when they had difficulty while alone, besides they could share their information. They could also finish their homework or task in a shorter time. On the other hand, it was also stated that some students did not like group work because there might be disagreement within the friends, and also they did not prefer group work because their ideas were not welcomed by their friends. The category of self-efficacy consisted of the code "suitability". The frequency values of this code also changed in favor of post-test. The codes in this category comprised of the thoughts that students found group work suitable for them, and realized that they could overcome this situation after the implementation, although they were biased and timid about group work with their friends. Some of the students' thoughts are given below:

S11: "...I sometimes like group work depending on the situation…"

S12: "...I prefer individual work because I think that the decision of a single person is more logical than a group decision…"

Table 9. Student views on the seventh question

Category	Code	Pre-test Frequency (F)	Post-test Frequency (F)	Pre-Test (%)	Post-test (%)
Yes	Usually	10	10	50	50
	Sometimes	4	4	20	20
	Rarely	2	2	10	10
No	Never	4	4	20	20
Total		20	20	100	100

In Table 9, the answers to the question "Can you take risks in situations that you think are important?" consisted of two categories called yes and no, and four codes (usually, sometimes, rarely, never). The category of yes consisted of codes that are called usually, sometimes and rarely. The frequency values of these codes remained stable. According to the statements of this category, students generally thought that they did not have any problems on risk taking, in other words, they did not refrain from taking risks as long as they could bear the consequences. Some of them sometimes took risks when the situation was important and some students rarely took risks. The category of no consisted of the code "never". The frequency values of this code also remained unchanged. The codes in this category comprised of the thoughts that students did not like to take risks at all or they were not inclined to take risks because it might be dangerous. So they never thought of taking risks. Some of the students' thoughts are given below:

S13: "...I wouldn't take any risks because it might be dangerous ..."

S14: "...I rarely take risks as it might cause bad consequences ..."

Conclusions and Suggestions

In this study, middle school students' entrepreneurial skills were investigated in terms of the entrepreneurship-based STEM education. Based on the results obtained from the quantitative data of this research, an increase was observed in favour of the post-test, although the overall increase was not statistically significant and also, according to the results obtained from the qualitative data, it was observed that the E-STEM activities contributed to the students' ability to work in groups, to stand out and leadership skills. In other words, the data obtained from the interviews indicated that E-STEM activities had a positive effect on the sub-dimensions of entrepreneurial skills called "leadership and tendency to stand out", "social skills and group work" and "the perception of innovation and creativity".

Many studies emphasize that one of the main goals of STEM education is to develop entrepreneurial skills (Guerra et al., 2014; O'Leary, 2012; Deveci, 2019; Caldwell et al., 2018; Rae & Melton, 2017; Camesano et al., 2016). Likewise, the quantitative and qualitative results of this research indicate that E-STEM education is a contributing factor that positively affects students' entrepreneurial skills. This result raises the question of which factors play a key role in the development of the entrepreneurial skills. Many studies have been trying to find an answer to the question of what factors have an impact on entrepreneurship. Stone et al. (2013) point out that among the factors affecting entrepreneurial tendency, cultural and ethnic origin plays a crucial role besides the social environment. In addition, this study clearly stated that the experimental studies in this field did not yield adequate results for a common idea or generalization. Tur-Porcar et al. (2018) have suggested based on their research findings that behavioral factors and ethical values are among the most important factors affecting sustainable entrepreneurship. According to these studies, it can be concluded that the development of entrepreneurial skills can not be affected only by an education programme or activity, genetic and environmental factors are also decisive in causing a sharp change. Considering the results of this study, the fact that the implementation period of the activities was limited to eight weeks may also be an impediment to a significant increase.

Some sub-dimensions and skills are crucial for the determination of entrepreneurial skills (Özcan, 2019). Considering the quantitative analysis on one of these sub-dimensions called "leadership and standing out", it was observed that the number of the students who liked leadership and standing out increased in favor of post-test. Students could take the opportunity to stand out and to act as a group leader and these experiences might have caused this behavioral change. The qualitative results support the quantitative analysis results of this sub-dimension. The fact that some of the students stated that they liked to thrust themselves forward in the group indicates that they used group work as a means to reveal their entrepreneurial characteristics. Some students who stated that they did not like to bring themselves to the forefront after the implementation or did not want to take organizing roles in group work changed their minds. Even partially, the activities had a positive effect on these skills of the students. The main reason for this result is that the students worked in groups during the activities and this process helped them discover leadership skills and reduce their prejudices. According to the analysis of quantitative data, although there is an increase in the leadership sub-dimension in favor of the post-test, this increase did not cause a significant change. Before the implementation, a significant portion of the students stated that leadership was not suitable for them or their organizational skills were not good, but after the implementation, some students stated that they believed that they could lead and organize the group. This result also partially supports the quantitative data. Therefore, it can be stated that the activities have a partially positive effect on the leadership tendencies of the students. The possible reason for this effect may be that leadership is influenced by experiences as well as personal characteristics. It can be claimed that students have gained an experience by leading the group, expressing their opinion, working collaboratively. Besides, this situation can be an indication that the concepts of leadership, standing out and entrepreneurship are interrelated.

Considering the quantitative analysis results of the sub-dimension called "the perception of innovation and creativity", an insignificant increase in favor of the post-test was observed. According to the answers obtained from the interview before the implementation, it was seen that most of the students could solve problems in at least one field. Among these fields, technical and social problems mostly stood out. After the implementation, some students gave examples of STEM-related problem solving in addition to the other fields. Students encountered problems during the applications of E-STEM activities and made an effort to solve the STEM-related problems. Therefore, in addition to the codes in the pre-test, during the post-test interview, the students stated that they could solve problems also in the field of STEM. At this point, it can be stated that the E-STEM activities contributed to the problem solving skills of some students. İnce et al. (2018) investigated the effect of STEM activities on children's problem solving skills and concluded that STEM activities had a positive effect on their children's problem solving skills. In addition, many studies have emphasized that STEM activities can positively affect problem solving and creativity skills also in a theoretical framework (Brown et al., 2011; Madden et al., 2013; Kelley & Knowles, 2016).

The concept of self-confidence which is another sub-dimension of entrepreneurial skills is described as the ability to express ideas and being decisive in the qualitative data. As a result, students' answers on being self-confident did not change significantly after the implementation. However, the answers given to the question about freely expressing their opinions positively increased. Students generally felt that their answers were wrong about solving the questions related to the course subjects. However, some students stated that they stood behind their decisions and trusted their own ideas. Based on this result, it can be inferred that the activities did not have a considerable effect on students' decisiveness in their opinions or decisions. Clifton and Gill (1994) have claimed that the concept of self-confidence is a psychological feature and they found supportive results in their study on the factors affecting self-confidence, also arguing that personal characteristics such as gender and perfectionism affect self-confidence.

According to the analysis of the thoughts about the fourth sub-dimension called "group work", it can be inferred that the qualitative and quantitative data results for this sub-dimension also support each other. It is also observed that there is a positive difference in students' views on group work. Many students who have prejudices about group work before the application of the activities, it was concluded that they actually enjoyed and managed to actively participate in the group work. It can be argued that the contribution of STEM activities in this result is obvious. The group studies experienced by the students within the scope of STEM activities positively changed the views of a significant portion of their students about group work. Likewise, in the studies conducted by Akdağ and Güneş (2017), Dönmez (2017), Akgündüz and Akpınar (2018), it was concluded that the STEM activities carried out with group studies contributed to the students' enjoyment of group work and collaboration.

There was an insignificant increase in favor of the post-test in the quantitative analysis of the final sub-dimension called "risk taking tendency". However, according to the qualitative data regarding risk taking, it was determined that the pre-and post-test frequencies of the answers did not change and the majority of the students were more or less risk-takers. Therefore, considering the qualitative results, STEM activities did not have an effect on students' risk taking tendency. On the other hand, according to the quantitative results, it can be claimed that the activities have an insignificant effect on this tendency. Bang (2000), has emphasized that the most important risk-taking factors are socio-cultural factors; such as age, gender, education level, and also in the study conducted by Saraç and Kahyaoğlu (2011) it has been revealed that socio-economic factors affect risk taking tendency. To sum up the results on these factors, similar studies indicate that risk-taking tendency is affected by many different factors and most of them are not caused by short-term environmental interventions. In fact, long-term variables and characteristics play a much more effective role on risk-taking behavior.

There may be many reasons why middle school students' entrepreneurial skills did not change significantly via STEM activities. Moreover, the studies pointed out that entrepreneurship is multidimensional, the factors affecting entrepreneurship are variable and numerous, the effect of the environmental factors, family and characteristics cannot be easily changed (Sharma & Madan, 2014; Kaygın & Güven, 2015; Fahed-Sreih et al., 2010). In addition, it can be asserted that the limited duration of the activities is a factor that prevents the change or development of entrepreneurial skills. Moreover, considering the pre-test data, the fact that the entrepreneurial characteristics of the participant group were not very low before the implementation is a factor that may cause this result. Consequently, the most comprehensive result that can be deduced that the STEM activities mostly affect the students' feelings and views on the concepts of group work, leadership, and standing out. Another result of the study is that activities did not play a significant role in characteristics and skills such as self-confidence and risk taking. Based on the increase in the post-test results of the process, which is limited to eight weeks, it can be predicted that this increase can be significant, provided that a much longer STEM application process is conducted. Based on this research, it is recommended to investigate the effects of other factors and educational approaches on entrepreneurial skills for future studies to be conducted on this subject. Besides, sub-dimensions of entrepreneurial skills can be discussed in more detail. The E-STEM implementation was carried out at the middle school level. However, this implementation can be replicated on students at other education levels. Finally, The implementation period can be kept longer to ensure a significant effect.

Acknowledgements

Ethics committee approval was obtained for this study (04/10/2019-No:10/03). This study has been produced from author's master dissertation.

Ortaokul öğrencilerinin girişimcilik becerilerinin Girişimcilik Temelli STEM Eğitimi Açısından İncelenmesi: Karma yöntem çalışması

Özet:

Bu çalışma ortaokul öğrencilerinin girişimcilik becerilerini E-STEM eğitimi açısından araştırmak amacıyla yapılmıştır. Araştırmanın evrenini kırsal kesimdeki ortaokul öğrencileri oluşturmaktadır. Tek grup ön ve son test deseninin kullanıldığı araştırmanın örneklemini 20 sekizinci sınıf öğrencisi oluşturmuştur. Araştırmada karma yöntem tercih edilmiştir. Nicel verileri elde etmek için girişimcilik ölçeği kullanılmıştır. Nitel verilerin elde edilmesi için yarı yapılandırılmış görüşme formu kullanılmıştır. Öğrencilere 8 hafta boyunca E-STEM eğitimi verilmiştir. Nicel verilerin ön ve son test ortalamaları bağımlı örneklem t testi ile karşılaştırılmış ve nitel verilerin analizinde içerik analizi yöntemi kullanılmıştır. Sonuç olarak, öğrencilerin girişimcilik becerileri ve girişimcilik alt boyutlarına ait puanları genel artış anlamlı olmasa da son test lehine yükselmiştir. Ayrıca nitel verilere göre E-STEM etkinliklerinin alt boyutlara olumlu etkisi olduğu görülmüştür.

Anahtar kelimeler: girişimcilik, ortaokul öğrencileri, STEM etkinlikleri

Anahtar kelimeler: 3 to 5 key words in Turkish, lowercase, comas between the key words.

About the Author(s)

Meryem Meral

A PhD student in Erzincan Binali Yıldırım University.
The main study fields: Science education, STEM, E-STEM, Educational diciplines
Mailing Address: Erzincan Binali YıldırımUniversity, Department of Science Education, Erzincan, Turkey, Zip,
Mobile: +90 543 388 95 73
E-mail: meryemmeral96@gmail.com

Sema Altun Yalçın

A proffessor in Erzincan Binali Yıldırım University.
The main study fields: Science education, STEM, E-STEM, Educational diciplines
Mailing Address: Erzincan Binali YıldırımUniversity, Department of Science Education, Erzincan, Turkey, Zip,
Mobile: +90 530 701 78 24
E-mail: saltun_11@hotmail.com

References

Abbasi, M. H., Siddiqi, A., & Azim, R. U. A. (2011). Role of effective communications for enhancing leadership and entrepreneurial skills in university students. *International Journal of Business and Social Science, 2*(10). doi: 10.30845/ijbss.

Ahmad, J., & Siew, N. M. (2021). Development of a Children Entrepreneurial Science Thinking Test for STEM Education. *Journal of Baltic Science Education*, *20*(4), 528-545. doi. 10.33225/jbse/21.20.528.

Akdağ, F. T., & Güneş, T. (2017). Determination of perceptions of Science High School students on energy and their levels of interdisciplinary association. *International Journal of Social Sciences and Education Research, 3*(5), 1643-1656. doi: 10.33225/jbse/21.20.52810.24289/ijsser.337238.

Akgündüz, D., & Akpınar, B. C. (2018). Evaluation of STEM Applications Based on Science Education in Pre-School Education in terms of Students, Teachers and Parents. *Yaşadıkça Eğitim*, *32* (1). doi: 10.17226/18290.

Ananiadou, K., & Claro, M. (2009). 21st Century Skills and Competences for New Millennium Learners in OECD Countries. doi: 10.1787/218525261154.

Avcı, Ö. (2018). *The effect of differential teaching applications on students' entrepreneurship skills and academic success* (Master's thesis). Kırıkkale University, Kırıkkale. Accessed from YÖK National Thesis Center.

Aydın, N.(2019). *The effect of STEM and STEM-based robotic activities on secondary school students' problem solving reflective thinking, mental risk taking and motivative strategies in learning* (Master's thesis). Erzincan Binali Yıldırım University, Erzincan. Accessed from YÖK National Thesis Center.

Aytaç, Ö. (2006). Entrepreneurship: A Socio-Cultural Perspective. *Dumlupınar Üniversitesi Sosyal Bilimler Dergisi*, (15).

Azoulay, P., Jones, B., Kim, J. D., & Miranda, J. (2018). Age and high-growth entrepreneurship (No. w24489). *National Bureau of Economic Research.*

Bang, Y. (2000). Factors affecting Korean students' risk-taking behavior in an EFL classroom. [Doctoral dissertation], *The Ohio State University.*

Briggs, C. L. (1986). Learning how to ask: A sociolinguistic appraisal of the role of the interview in social science research (No. 1). *Cambridge University Press*.

Brown, R., Brown, J., Reardon, K., & Merrill, C. (2011). Understanding STEM: current perceptions. *Technology and Engineering Teacher, 70*(6), 5. doi: 10.1136/bjsports-2011-090606.55.

Büyüköztürk, Ş., Kılıç Çakmak, E., Akgün, Ö. E., Karadeniz, Ş. ve Demirel, F. (2016). *Bilimsel Araştırma Yöntemleri [Scientific Research Methods]*. (21. Baskı), *Pegem Akademi*: Ankara.

Caldwell, L., Garcia, R., & Cagle, N. (2018). K-12 diversity pathway programs in the E-STEM fields: A review of existing programs and summary of unmet needs. *Journal of STEM Education: Innovations and Research, 19*(4), 12-18.

Camesano, T. A., Billiar, K., Gaudette, G., Hoy, F., & Rolle, M. (2016). Entrepreneurial mindset in STEM education: Student success. In *VentureWell. Proceedings of Open, the Annual Conference* (p. 1). National Collegiate Inventors & Innovators Alliance.

Caruth, G. D. (2013). Demystifying mixed methods research design: A review of the literature. *Online Submission*, *3*(2), 112-122. doi: 10.13054/mije.13.35.3.2.

Clifton, R. T., & Gill, D. L. (1994). Gender differences in self-confidence on a feminine-typed task. *Journal of sport and exercise psychology*, *16*(2), 150-162. doi: 10.1123/jsep.16.2.150.

Creswell, J. W., & Miller, D. L. (2000). Determining validity in qualitative inquiry. *Theory into practice, 39*(3), 124-130. doi: 10.1207/s15430421tip3903_2.

Creswell, J. W. (2013). Steps in conducting a scholarly mixed methods study.

Çakır, Z. (2018). *Investigation of the effects of montessori approach based STEM effectiveness on pre-school teacher candidates* (Master's thesis). Erzincan Binali Yıldırım University, Erzincan. Accessed from YÖK National Thesis Center.

Daniel, A. D., Costa, R. A., Pita, M., & Costa, C. (2017). Tourism Education: What about entrepreneurial skills?. *Journal of Hospitality and Tourism Management, 30*, 65-72. doi: 10.1016/j.jhtm.2017.01.002.

Deveci, İ. (2016). *The development, implementation and evaluation of entrepreneurship training modules integrated with the middle school science curriculum* (Doctoral thesis). Uludağ University, Bursa. Accessed from YÖK National Thesis Center.

Deveci, İ., Zengin, M. N.,& Çepni, S. (2015). Development and evaluation of science based enterprising training module. *Journal of Educational Sciences & Practices*, *14*(27).

Deveci, İ. (2018). The STEM Awareness as Predictor of Entrepreneurial Characteristics of Prospective Science Teachers. *Kastamonu Eğitim Dergisi*, *26*(4), 1247-1256. doi: 10.24106/kefdergi.356829.

Deveci, İ. (2019). Reflections with regard to Entrepreneurial Project (E-STEM) Process on the Life Skills of Prospective Science Teachers: A Qualitative Study.

Dönmez, İ. (2017). The Views of Students and Team Coaches about Robotic Competitions on the STEM Education Framework (Case of First Lego League). *Eğitim Bilim ve Teknoloji Araştırmaları Dergisi, 2*(1), 25-42.

Etikan, I., Musa, S. A., & Alkassim, R. S. (2016). Comparison of convenience sampling and purposive sampling. *American journal of theoretical and applied statistics, 5*(1), 1-4. doi: 10.11648/j.ajtas.20160501.11.

Ezeudu, F. O., Ofoegbu, T. O., & Anyaegbunnam, N. J. (2013). Restructuring STM(Science, Technology, and Mathematics) Education for Entrepreneurship. *Online Submission, 3*(1), 27-32.

Fahed-Sreih, J., Pistrui, D., Huang, W. V., & Welsch, H. P. (2010). Family and cultural factors impacting entrepreneurship in war time Lebanon. *International Journal of Entrepreneurship and Innovation Management*, *12*(1), 35-51. doi: 10.1504/IJEIM.2010.033166.

Farwati, R., Metafisika, K., Sari, I., Sitinjak, D. S., Solikha, D. F., & Solfarina, S. (2021). STEM education implementation in Indonesia: a scoping review. *International Journal of STEM Education for Sustainability*, *1*(1), 11-32. doi.10.53889/ijses.v1i1.2.

Gonzalez, H. B., & Kuenzi, J. J. (2012). Science, technology, engineering, and mathematics (STEM) education: A primer. Washington, DC. *Congressional Research Service,* Library of Congress.

Guerra, R. C. C., Smith, K. A., McKenna, A. F., Swan, C., Korte, R., Jordan, S., ... & MacNeal, R. (2014, October). Innovation corps for learning: Evidence-based entrepreneurship™ to improve (STEM) education. In *2014 IEEE Frontiers in Education Conference (FIE) Proceedings* (pp. 1-5). IEEE.

Hall, J. N. (2013). Pragmatism, evidence, and mixed methods evaluation. *New Directions for Evaluation*, *2013* (138), 15-26. doi: 10.1002/ev.20054.

Hindle, K., & Rushworth, S. (2000). Yellow pages global entrepreneurship monitor (gem) australia 2000. *University of Illinois at Urbana-Champaign's Academy for Entrepreneurial Leadership Historical Research Reference in Entrepreneurship.*

Hsieh, H. F., & Shannon, S. E. (2005). Three approaches to qualitative content analysis. *Qualitative health research, 15*(9), 1277-1288. doi: 0.1177/1049732305276687.

Ivankova, N. V., Creswell, J. W., & Stick, S. L. (2006). Using mixed-methods sequential explanatory design: From theory to practice. *Field methods, 18*(1), 3-20. doi: 10.1177/1525822X05282260.

İnce, K., Mısır, M. E., Küpeli, M. A., & Fırat, A. (2018). Examining the effect of STEM-Based Approach on the Problem Solving Ability and Academic Success of Students in Teaching the Enigma of the Earth's Crust Unit of the 5th Grade Life Sciences Course. *Journal of STEAM Education*, *1*(1), 64-78.

Johnson, P. (2009). The 21st century skills movement. *Educational Leadership, 67*(1), 11.

Kaygın, E. & Güven, B. (2015). *Güçlü Kadınlar: Türkiye'de kadının girişimciliği.[Strong women:The entrepreneurship of woman in Turkey]*. *Veritaş akademi.*

Kelley, T. R., & Knowles, J.G. (2016). A conceptual framework for integrated STEM education. *International Journal of STEM Education, 3*(1), 11. doi: 10.1186/s40594-016-0046-z.

Konuş, F.Z. (2019). *Entrepreneurial tendency as predictor of stem attitudes of seventh and eighth grade students* (Master's thesis). Sütçü İmam University, Kahramanmaraş. Accessed from YÖK National Thesis Center.

Langdon, D., McKittrick, G., Beede, D., Khan, B., & Doms, M. (2011). STEM: Good Jobs Now and for the Future. *ESA Issue Brief, 03-11. US Department of Commerce.*

Leech, N. L., Dellinger, A. B., Brannagan, K. B., & Tanaka, H. (2010). Evaluating mixed research studies: A mixed methods approach. *Journal of mixed methods research, 4*(1), 17-31. doi: 10.1177/1558689809345262.

Lepuschitz, W., Koppensteiner, G., Leeb-Bracher, U., Hollnsteiner, K., & Merdan, M. (2018). Educational Practices for Improvement of Entrepreneurial Skills at Secondary School Level. doi: 10.3991/ijep.v8i2.8141.

Lincoln, Y. S., & Guba, E. G. (2000). The only generalization is: There is no generalization. *Case study method*, 27-44.

Madden, M. E., Baxter, M., Beauchamp, H., Bouchard, K., Habermas, D., Huff, M., ... & Plague, G. (2013). Rethinking STEM education: An interdisciplinary STEAM curriculum. *Procedia Computer Science, 20,* 541-546. doi: 10.1016/j.procs.2013.09.316.

Maxwell, J. (1992). Understanding and validity in qualitative research. *Harvard educational review, 62*(3), 279-301. doi: 10.17763/haer.62.3.8323320856251826.

Mee, R. W., & Chua, T. C. (1991). Regression toward the mean and the paired sample t test. *The American Statistician, 45*(1), 39-42. doi: 10.1080/00031305.1991.10475763.

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

Morse, J. M., & Cheek, J. (2015). Introducing qualitatively-driven mixed-method designs. *Researrch Methods & Evaluationt*, *25* (6), 731-733. doi: https://doi.org/10.1177/1049732315583299.

Nafukho, F., Kobia, M., & Sikalieh, D. (2010). Towards a search for the meaning of entrepreneurship. *Journal of European industrial training*, *34* (2), 110-127. doi: 10.1108/03090591011023970.

Nartgün, Ş., & Kaya, A.(2016). Creating school image in accordance with private school parents' expectations. *Journal of Research in Education and Teaching*, *5* (2), 154-167.

Nicolaides, A. (2011). Entrepreneurship-the role of higher education in South Africa. *Educational research*, *2*(4), 1043-1050.

O'Leary, S. (2012). Impact of entrepreneurship teaching in higher education on the employability of scientists and engineers. *Industry and Higher Education*, *26*(6), 431-442. doi: 10.5367/ihe.2012.0128.

Özcan, E. (2019). *Effects of problem based learning on prospective science teachers' problem solving skills, academic achievements and attitudes* (Doctoral thesis). Dokuz Eylül University, İzmir. Accessed from YÖK National Thesis Center.

Özkızılcık, M., & Cebesoy, Ü. B. (2020). The Effectiveness of Design Based STEM activities on Pre-service Science Teachers' Problem Solving Skills and STEM Teaching Intentions. *Uludağ Üniversitesi Eğitim Fakültesi Dergisi*, *33*(1), 177-204. doi: 10.19171/uefad.588222.

Öztürk, S.C.(2018). *The effects of STEM education on the problem solving and critical thinking skills of science teachers* (Master's thesis). Erzincan Binali Yıldırım University, Erzincan. Accessed from YÖK National Thesis Center.

Rae, D., & Melton, D. E. (2017). Developing an entrepreneurial mindset in US engineering education: an international view of the KEEN project. *The Journal of Engineering Entrepreneurship*, *7*(3).

Razali, N. M., & Wah, Y. B. (2011). Power comparisons of shapiro-wilk, kolmogorov-smirnov, lilliefors and anderson-darling tests. *Journal of statistical modeling and analytics*, *2*(1), 21-33.

Rocha, H., & Birkinshaw, J. (2007). Entrepreneurship Safari (Vol. 12). *Now Publishers Inc.*

Ross, S. M., & Morrison, G. R. (2004). Experimental research methods. *Handbook of research on educational communications and technology, 2*, 1021-43.

SA, N. H., Arifi, A., Putranta, H., & Azizah, A. N. M. (2021). Experiences of participants in teacher professional education on obtaining soft skills: A Case Study in Indonesia. *European Journal of Educational Research, 10*(1), 313-325. doi: 10.12973/eu-jer.10.1.313.

Sánchez, J. C. (2013). The impact of an entrepreneurship education program on entrepreneurial competencies and intention. *Journal of small business management, 51*(3), 447-465. doi: 10.1111/jsbm.12025.

Saraç, M., & Kahyaoğlu, M. B. (2011). The analysis of socio-economic and demographic factors effecting the risk taking behaviour of individual investors. *Journal of BRSA Banking & Financial Markets,5*(2).

Shahin, M., Ilic, O., Gonsalvez, C., & Whittle, J. (2021). The impact of a STEM-based entrepreneurship program on the entrepreneurial intention of secondary school female students. *International Entrepreneurship and Management Journal*, *17*(4), 1867-1898. doi. 10.1007/s11365-020-00713-7.

Stone, D. L., Canedo, J. C., Harrison, T. L., Lukaszewski, K. M., Suazo, M., & Krueger, D. C. (2017). The Relations Between Entrepreneurs' Ethnicity, Familism, Values, Beliefs, and Use of Financial Planing. *Journal of Business and Entrepreneurship*, *28*(2), 50-81.

Şencan, H. (2005). *Güvenilirlik ve geçerlilik.[Reliability and validity]*. Hüner Şencan.

Sharma, L., & Madan, P. (2014). Effect of individual factors on youth entrepreneurship–a study of Uttarakhand state, India. *Journal of Global Entrepreneurship Research, 4(*1), 3. doi: 10.1186/2251-7316-2-3.

Stevenson, H. H. (2000). Why entrepreneurship has won. *Coleman White Paper*, *2*(4), 483.

Tarhan, M., & Kılıç, F. (2017). Bibliography of resources for improving entrepreneurship skills in social studies education. *Abant İzzet Baysal Üniversitesi Eğitim Fakültesi Dergisi*. doi: 10.17240/aibuefd.2017.17.32772-363986.

Taşpınar, M. (2017). *Sosyal Bilimlerde SSPS Uygulamalı Nicel Veri Analizi.[SPSS Applied Quantitative Data Analysis in Social Sciences]*.(1. Baskı), *Pegem Akademi:* Ankara.

Topsakal, İ.(2018). *The research of the system education on the student's learning environment, critical thinking disposition and problem solving abilities perception*

(Master's thesis). Erzincan Binali Yıldırım University, Erzincan. Accessed from YÖK National Thesis Center.

Tur-Porcar, A., Roig-Tierno, N., & Llorca Mestre, A. (2018). Factors affecting entrepreneurship and business sustainability. *Sustainability*, *10*(2), 452. DOİ: 10.3390/su10020452.

Trilling, B., & Fadel, C. (2009). 21st century skills: Learning for life in our times. John Wiley & Sons.

Türnüklü, A. (2000). Eğitimbilim araştırmalarında etkin olarak kullanılabilecek nitel bir araştırma tekniği: Görüşme. *[A qualitative research technique that can be used effectively in pedagogical research: Interview]. Kuram ve Uygulamada Egitim Yönetimi Dergisi*, *6*(4), 543-559. Retrieved from https://dergipark.org.tr/tr/pub/kuey/issue/10372/126941.

Uygun, M., & Güner, E. (2016). The Role of Entrepreneurship Education in Developing Students' Entrepreneurial Intentions. *Manas Sosyal Araştırmalar Dergisi*, *5*(5), 37-57. Retrieved from https://dergipark.org.tr/en/pub/mjss/issue/40508/485490.

Venkatesh, V., Brown, S. A., & Bala, H. (2013). Bridging the qualitative-quantitative divide: Guidelines for conducting mixed methods research in information systems. *MIS quarterly*, 21-54. doi: 10.25300/MISQ/2013/37.1.02.

Walan, S. (2019). The dream performance–a case study of young girls' development of interest in STEM and 21st century skills, when activities in a makerspace were combined with drama. *Research in Science & Technological Education*, 1-21. doi: 10.1080/02635143.2019.1647157.

Necatibey Eğitim Fakültesi Elektronik Fen ve Matematik Eğitimi Dergisi
Cilt 16, Sayı 2, Aralık 2022, sayfa 354-380. ISSN: 1307-6086
Necatibey Faculty of Education Electronic Journal of Science and Mathematics Education
Vol. 16, Issue 2, December 2022, pp. 354-380. ISSN: 1307-6086

Research Article

Development of Basic Biotechnology Knowledge Scale (BBKS) with Rasch Measurement Model

Serap OZ AYDIN[1], Mesut SACKES [2], Nazlı Ruya TASKIN BEDIZEL[3], Aysun SICAKER[4]

[1] Balikesir University, Necatibey Faculty of Education, Balikesir, soz@balikesir.edu.tr, https://orcid.org/0000-0002-0635-0728

[2] Balikesir University, Necatibey Faculty of Education, Balikesir, msackes@gmail.com, https://orcid.org/0000-0003-3673-1668

[3]Balikesir University, Necatibey Faculty of Education, Balikesir, nazliruya@balikesir.edu.tr, https://orcid.org/0000-0001-6027-719X

[4]Manisa Soma İstanbul Science High School, Manisa, asicaker@hotmail.com, https://orcid.org/0000-0001-7615-7440

Received : 12.12.2022 Accepted : 29.12.2022

Doi: 10.17522/balikesirnef.1217928

Abstract – The purpose of this study was to develop a scale to measure basic biotechnology knowledge, examine the psychometric properties of the scale, and investigate whether there are differences in the test performances of high school students in terms of school, grade, and gender. The development phase of this scale was carried out with a sample of 388 high school students in a province in the west of Turkey. The psychometric properties of the scale were examined using the Rasch model. The K-R internal consistency coefficient of the final scale consisting of 17 items was calculated as 0.77. It was observed that item-total correlations varied between 0.25 and 0.48 except for one item (item 1, 0.13). The results of the Rasch analysis indicated that the scale fits the Rasch model and can differentiate between low and high-performing test takers. Three-Way ANOVA results demonstrated a significant main effect for the school. There were no statistically significant differences for grade and gender variables in terms of their biotechnology knowledge scores. However, the grade*gender interaction was statistically significant, favouring males with a small effect size. This observed effect was possibly due to the uneven sample size of 12th-grade students. The overall results suggest that Basic Biotechnology Knowledge Scale (BBKS) can be used to assess the biotechnology knowledge level of high school students.

Keywords: biotechnology, knowledge measurement, Rasch measurement model, scale development

----

Corresponding author: Nazli Ruya TASKIN BEDIZEL, nazliruya@balikesir.edu.tr, Balikesir University

Introduction

Biotechnology is a branch of science that arose from scientific curiosity and is critical for students to comprehend because of its possible influence on them and others (Kustiana, Suratno, & Wahyuni, 2020). By providing a wide range of products, rapid developments in biotechnology and genetic engineering have shown their impact on our lives directly or indirectly in many areas, such as health, agriculture, the environment, and food production (Ayar & Hasipek, 2003; Lyson, 2002; Özgen, 1995). However, not knowing the future results of the biotechnological developments (Ho, 2001) caused biotechnological applications to be accepted as risky, particularly in areas such as health and the environment (Kahveci & Özçelik, 2008; Shaw, 2002). In general terms, medical procedures, studies of microorganisms, plants, and environmental studies are perceived positively, but human cloning, direct human work, and non-compulsory and more arbitrary (such as making food more caloric) applications are viewed negatively (Akman, 2007; Bayoğlu & Özgen, 2010; Chabalengula, Mumba, & Chitiyo, 2011; Demir & Pala, 2007; Frewer, Shepherd, & Sparks, 1994; Frewer, Howard, & Shepherd, 1997; Gardner & Jones, 2011; Gardner, Jones, Taylor, Forrester, & Robertson, 2010; Morris & Adley, 2001; Schilling, Hallman, Hossain, & Adelaja, 2003). Studies also indicate that individuals have gained some of their knowledge through informal means such as TV and newspapers and usually have low-level and simple information about biotechnology (Gaskell et al., 2006; Sjöberg, 1996, Sparks & Shepherd, 1994). For example, Bonfadelli (2005) states that the amount of biotechnology information covered in the media is directly proportional to the knowledge of biotechnology. Therefore, the information individuals obtain informally may not be of the nature to raise awareness, and in this case, they need structured ways of learning. This requires societies of conscious individuals in the field of biotechnology (Harms, 2002).

Teaching biotechnology topics in schools can help students become 'biotechnologically literate people' who grasp both the concepts of current biotechnology and the fundamental principles of biotechnology. This provides opportunities for them to build views and consequences of biotechnology that will allow them to make educated personal and social decisions (Gonzalez, Casanoves, Salvado, Barnett, & Novo, 2013; Paš, Vogrinc, Raspor, Kneževič, & Zajc, 2019). However, despite its significance and rapid development, studies (e.g. Fonseca, Costa, Lencastre, & Tavares, 2012) indicate that individuals generally have poor knowledge of biotechnology. It has not been a popular topic, particularly in public schools, due to teachers' inadequate academic skills, limited time and to the lack of resources available (Fonseca et al., 2012; Gelamdin, Alias, & Attaran, 2013).

The Assessment of Biotechnology Knowledge

In Turkey, from elementary school to post-secondary education, science courses include the multidisciplinary area of biotechnology (MoNE, 2018a, b). In the Primary Science Course Curriculum, biotechnology topics are included in the 8th grade 2nd Unit called "DNA and Genetic Code" (MoNE, 2018a). In the High School Biology Course Curriculum, they are included in the 12th grade 1st unit called "From Gene to Proteins" (MoNE, 2018b). In the high school Genes to Proteins Unit, Gene Technologies, DNA Fingerprint, Stem Cell Technologies, Model Organisms, Genetic Consulting, Cloning, Gene Therapy Applications, Vaccines, Bioethics and Biosecurity topics are covered (MoNE, 2018b). One of the observable ways in which the applied programs accomplish their goals is to evaluate how well the knowledge and competencies of the students following this curriculum improve in line with their abilities. However, the studies on the knowledge of genetic engineering and biotechnology are diverse in terms of various factors such as the target audience, the type of data collection tools used, the breadth of the data collection tool and data analysis methods. To examine if individuals have the basic knowledge of biotechnology, accurate and efficient biotechnology knowledge measurement is needed. Over the past twenty-five years, numerous scales for measuring different aspects of biotechnology have been developed worldwide. When the studies are examined, it is seen that a substantial number of studies are about the attitudes towards biotechnology (Bal & Keskin, 2002; Bilen & Özel, 2012; Massarani & Moreira, 2005; Sürmeli & Şahin, 2010a,b; Turan & Koç, 2012) and knowledge of genetic engineering and biotechnology (Acarlı, 2016; Ağaç, 2019; Akman, 2007; Chen & Raffan, 1999; Dawson, 2007; Dawson & Schibeci, 2003; Gürkan & Kahraman, 2019; Konak & Hasancebi, 2021; Keskin et al., 2010; Prokop et al., 2007; Sıcaker & Öz Aydın, 2015; Sıcaker, Öz Aydın, & Saçkes, 2020; Sönmez & Pektaş, 2017; Yüce & Yalçın, 2012).

There are various measurement tools using different types of questions to evaluate students' knowledge of biotechnology. Examples of these are open-ended questions (Chen & Raffan, 1999; Dawson & Schibeci, 2003; Kinderlerer & Beyleveld, 1998; Lock & Miles, 1993), true-false questions (Casanoves, González, Salvadó, Haro, & Novo, 2015; de la Hoz, Solé-Llussà, Haro, Gericke, & Valls, 2022; Gürkan & Kahraman, 2019; Klop & Severiens, 2007; Prokop, Leskova, Kubiatko, & Diran, 2007; Sıcaker, Öz Aydın, & Saçkes, 2020), Likert type scales (Lamanauskas & Makarskaitė-Petkevičienė, 2008; Yüce & Yalçın, 2012) and multiple choice questions (Atasoy, Atıcı, Taşar, & Taflı, 2020) tested on various samples including university students, pre-service teachers, elementary school and secondary school students. It seems from the measurement tools that there is a wide variety of measures in different countries for different samples and covering various topics of biotechnology. When the topics, samples and contexts questioned in biotechnology knowledge studies conducted in various countries are examined, it is seen that the instruments cover a wide range of topics such as the meaning of biotechnology and related concepts with examples, food biotechnology, cloning, genetically modified organisms, animal reproduction, animal reproduction, bioremediation, biotechnology ethics, electrophoresis, environmental and microbial biotechnology, gene splicing, growth hormones, hybridization, human genomics, plant-tissue culture, recombinant DNA, resistant plant species, transgenic species and various applications of biotechnology (Chen & Raffan, 1999; Lock & Miles, 1993; Mowen, Roberts, Wingenbach, & Harlin, 2006; Priest, Bonfadelli, & Rusanen, 2003; Prokop et al., 2007; Sıcaker & Öz Aydın, 2015).

This above-mentioned diversity in the studies calls for a need to design a tool that focuses on both the Turkish elementary science and high school Biology curriculum and on the other areas needed to ensure biotechnological literacy. In this case, issues such as validity, reliability and statistical methods become much more important. While it is very difficult to fully provide these with classical test methods (Boone & Scantlebury, 2006) as in the many existing scales guided, one of the item response theories, the Rasch measurement model, is one of the methods recommended to solve all these problems (Wright & Mok, 2004).

Recently, several authors (Sıcaker et al., 2020) have expressed the need for different measurement approaches, such as Item Response Theory (IRT) and Rasch analysis. Rasch measurement model has some advantages in developing and validating scales investigating the extent to which an item set meets several criteria essential for accurate measurement (Woudstra et al., 2019). First, it helps researchers to make critical corrections while using raw test score data allowing nonlinear raw data to be converted to a linear scale (Boone, 2016). Second, it provides the opportunity to evaluate the individuals according to their abilities and the items according to their difficulties.

In addition, Rasch Measurement Model evaluates every individual independently from the sample (Boone & Scantlebury, 2005; Wright & Mok, 2004). There are Rasch steps that may be employed to investigate more significant instrumentation issues such as item reliability, person reliability, and differential item functioning (Boone, 2016). All these advantages indicate that Rasch models can be easily used for two-category scales, such as True/False and Yes/No (), by overcoming the chance factor (Boone & Scantlebury, 2006; Wright & Mok, 2004). Recently, there seems to be a growing interest in using Rasch analysis in the science education field for scale development (Saefi et al., 2020; Testa et al., 2022; Tyas, Senam, Wiyarsi, & Laksono, 2020). Experimental studies and scale development studies using the Rasch model are also available in the fields of medicine and educational sciences (Baştürk, 2010; Elhan & Atakurt, 2005; Kaptan, 1994; Kaskatı, 2011; Koparan & Güvenen, 2013; Semerci, 2011a, b). Almost all these studies show that the use of Rasch models leads to better and more effective outcomes in evaluation and assessment (Sıcaker, 2013).

The Present Study

The purpose of the present study is to develop a Basic Biotechnology Knowledge Scale (BBKS) using Rasch Measurement Model and examine whether there is a statistical difference in biotechnology knowledge among students in terms of school, grade, and gender. The specific research questions for the present study are:

(1) What is the evidence to suggest the validity and reliability of measures of the Biotechnology Knowledge Scale?

(2) Is there a statistically significant difference in biotechnology knowledge between female and male students among students of different high school grades and different high schools?

Method

Study Model

The present study is a scale development study structured based on a survey model. Usually, at a specific point in time, surveys collect data to explain the existing conditions, define criteria against which existing conditions can be measured, or assess the relationships that occur between events. Surveys are also useful in generating accurate instruments through piloting and revision (Cohen, Manion, & Morrison, 2018).

Development Process of the Scale

The following steps were used to develop the BBKS: Conceptualizing the construct; creating the initial item pool; evaluating and modifying the items; conducting cognitive interviews; developing the pilot test; and validating the scale. A Three-Way ANOVA was used to analyze the differences in biotechnology knowledge among students from various schools, genders, and grade levels. For collecting valid evidence for BBKS to answer the first research question, AERA (American Educational Research Association), APA (American Psychological Association) and NCME (National Council on Measurement in Education) 2014 standards and guidelines were referred.

Conceptualizing the Scale

The advancements in biotechnology have had a significant positive impact on society and modern science. Nevertheless, despite the importance of the area, biotechnology receives scant attention in curricula and classrooms (Borgerding, Sadler, & Koroly, 2013; Hanegan & Bigler, 2009). For example, In Turkey, biotechnology topics are only covered in 12th grade, and despite the mentioned importance, only students who are enrolled in science-based courses encounter this subject. This brings the situation to the point that other students only encounter biotechnology subjects in 8th grade, only for four hours and with limited outcomes (MoNE, 2018a). Considering the effect of learning about biotechnology and resulting skills on students' interest and motivation in science (Hanegan & Bigler, 2009; Nordqvist & Aronsson, 2019), the fact that biotechnology is included in the programs so narrow and that not all students encounter these subjects sufficiently guided the development. Therefore, in the conceptualization of BBKS, attention was paid to include both elementary and high school biotechnology outcomes and other current developments in the biotechnological field.

Creating the Initial Item Pool

To generate the scale items, the first Turkish High School Biology Curriculum and curriculum-related textbooks were reviewed. Also, not limited to the program alone, some items, including current biotechnology topics that are thought to be known by all high school students, were also added to the scale. Based on the first review, the fourth author's discussions with her students and the researchers' experience, 37 short answer and true/false questions were prepared for the biotechnology and genetic knowledge of high school students. These 37 questions were informally tested on high school students in the fourth author's classrooms. In the second step, 84 items in True-False format were prepared as an initial item pool according to the results of multiple-choice questions. To examine items with lower and higher content validity, three experts in biology education with more than 20 years of experience were invited to review the initial item pool and asked to evaluate each item in the initial scale if the item is suitable to measure the biotechnology knowledge. First, opinions were received from field experts to gather evidence based on test content. In this initial review, items such as "*DNA can be completely cloned out of Vivo by PCR method*" and "*Methods such as mutation and crossing-over are the biotechnological methods of nature*" were excluded according to experts' opinions since they stated there were similar and more suitable items measuring the same content.

Conducting Cognitive Interviews

To provide construct-related validity evidence, think-aloud procedures were conducted with 15 high school students and two master's students in biology education. The cognitive interview process is an iterative process in scale development research to revise the content with one-on-one interviews (Willis, 2005). This interview process helped the researchers to that the items in BBKS were interpreted in the way that it is intended to measure and that the selected options of students reflected their thoughts. After this think-aloud procedure, some items were excluded, and the 84-item scale became a four main-topic, 16 sub-topics, 44-item scale named Basic Biotechnology Knowledge Scale (BBKS).

Developing the Pilot Test-First Application

Rasch analysis assumes that the probability of a person choosing a category of any item is a logistic function of the difference between the level of ability of the person and the level of difficulty of the item (Koparan & Güvenen, 2013). There are some problems encountered when trying to evaluate any questionnaire or test using the raw scores obtained by collecting the correct answers given to the items. One of them is the inability to determine the unexpected answers given to the items, that is, an item that is answered correctly by chance (especially in True/False tests). It is not possible to predict whether the correct answer was given knowingly or unknowingly. The Rasch measurement model has a structure that can overcome these problems (Wright & Mok, 2004). In this study, the following assumptions are examined to check the suitability of the Rasch analysis for sampling: Examining fit indices (item reduction), unidimensionality, local independence of items, person raw score reliability, separation indices, analysis of biased items, and examining the Wright Item-person map. WINSTEPS 3.65.0 is used to analyze the data with the Rasch Measurement Model.

For the pilot study, the 43-item version of BBKS was conducted on a sample of 150 11th and 12th-grade students from two public high schools in a province in the west of Turkey. The data were analyzed using Rasch analysis. These 43 items were grouped under four main topics: Basic Knowledge (12 items), Real-Life Practices (8 items), Laboratory Methods Techniques (14 items), and Effects (8 items). The evaluation of the Wald test for item elimination and item (category) difficulty parameters (Beta) results showed that 12 items were not fit the model. The researchers decided to keep five of these 12 items since the content validity is affected by their elimination. After reviewing the five items and excluding seven unsuitable items, the pilot analysis resulted in having 36 items in BBKS. The K-R20 internal consistency coefficient of this version of BBKS consisting of 36 items was calculated as 0.70. The five items that were edited and added to the BBKS and the seven items that were removed are presented in Table 1.

Table 1. Excluded and reviewed items after the pilot analysis

Item Numbers	Item Statement	Reviewed statement of the item
Item 1	Gene transfer cannot be made between organisms that are genetically quite different from each other (such as bacteria and humans).	Gene transfer can be made between living things (such as bacteria and humans) that are genetically quite different from each other.
Item 8	Gene therapy is a very easy and effective method	Gene therapy in humans is an easily applicable method
Item 10	Plants cannot produce animal proteins, even with genetic changes	By transferring genes to plants, they can be made to produce animal proteins.
Item 13	Cloning studies are not applicable to plants.	Excluded
Item 17	Biotechnological methods can only be applied in the laboratory	Excluded
Item 18	Genetic engineering only works on animal organisms	Excluded
Item 21	Stem cells are not found in all multicellular organisms	Excluded
Item 25	Gene (DNA) transfer to all plants occurs only through soil bacteria	Excluded
Item 28	Developing DNA technology does not pose significant ethical problems	Developing DNA technology may pose significant ethical problems
Item 31	Humans have fewer genes than most plants and invertebrates	Excluded
Item 33	Eggs and sperm of mammals cannot be combined outside of a living thing	Excluded
Item 35	DNA cannot be replicated outside the cell; in the laboratory	DNA cannot be replicated outside the living cell (under laboratory conditions)

Validating the Scale-Second Application

To provide valid evidence based on internal structure, the 36-item version of BBKS was subsequently applied to 388 high school students enrolled in various high schools in a province in the west of Turkey, and Rasch analysis tested the psychometric properties of the scale items. Of this sample, 209 were female (53.86%), and 179 were male (46.14%). HS1 and HS2 are Anatolian High Schools, and HS3 is a Science High School. The difference between HS1 and HS2 is the high school acceptance scores of students, which is higher for HS1 than HS2. Also, HS3 is a science-intensive high school, and its acceptance scores are higher than the other two. Table 2 shows the demographics of the sample.

Table 2. Demographics of the second application sample

Variables				N	(%)
School	HS1	10th grade	59	165	42.53
		11th.grade	59		
		12th grade	47		
	HS2	10th grade	59	77	19.84
		11th grade	18		
		12th grade	0		
	HS3	10th grade	119	146	37.63
		11th.grade	27		
		12th grade	0		
Grade	10th grade			237	61.08
	11th.grade			104	26.81
	12th grade			47	12.11
Gender	Female			209	53.86
	Male			179	46.14
Total				388	100

The difficulty of each item (β) was calculated, and items that did not fit the Rasch model were determined and excluded from the scale by examining the Wald test results, and the analyses were repeated. Nineteen items were excluded from the scale as a result of eliminating the items incompatible with the model. As a result of the second application analysis, 17 items were identified in the final version of the scale three main topics emerged: (1) Laboratory Methods and Techniques (Items 9, 13, 14, 19, 30, 36), (2) Real Life Practices (Items 1, 7, 18, 28, 33), and (3) Effects of Biotechnology (Items 2, 10, 16, 29, 31, 35)

Analysis of the Variance

In order to test the second research question to examine the practicability of BBKS, a three-way ANOVA was performed to compare the differences in biotechnology knowledge using three levels of school, three levels of grade (10, 11 and 12) and two levels of students' gender (boys and girls) to examine school, grade, and gender as between-subject factors. Inspection of the test assumption suggested no major deviations. Analyses were performed using SPSS version 26.

Findings

Psychometric Properties of the Items in BBKS

Here, the difficulty of each item (β) was calculated with Rasch analysis, and by examining Wald test results, items that did not fit the Rasch model were determined and removed from the scale, and the analyzes were repeated. Nineteen items incompatible with the model were excluded from the scale. The K-R20 internal consistency coefficient of the final scale consisting of 17 items was calculated as 0.77. It was observed that item-total correlations varied between 0.25 and 0.48 except for one item (Item 1, 0.13).

Rasch Analysis Findings of the Items in BBKS

Table 3 presents the item difficulty (b), or location, parameters for the 17-item BBKS. Item difficulty (b) shows where the item functions best along the trait scale. When the b value is lower, it means the item is "easier and expected to be endorsed at lower trait levels." (Nguyen, Han, Kim, & Chan, 2014, p.3). The item with a value of zero is of medium difficulty, and the item's difficulty level increases as it moves away from zero in the (+) direction, and its ease level increases as it moves away from zero in the (-) direction. As also shown in Table 4, according to Rasch's analysis, the most difficult item in the scale is "Item 13", and the easiest item is "Item 7". Also, "Item 31" is closest to medium difficulty. When the item map is examined in general, it is seen that the distribution of the number of easy and difficult items in the scale is equal. When it is examined according to Item 31, and it is accepted that the scale is close to medium difficulty, It is seen that eight items are more difficult than medium level and eight items are less difficult than medium level. The fact that each of the items is at different levels indicates that the scale has a homogeneous distribution in terms of item difficulties.

Table 3. Item Difficulty Parameters (b) and Item Statements in BBKS

Item Numbers	Item Statements	Est (b) (Logit)	Std. Error	Lower CI.	Upper CI.	Item- Total Correlation
Item 13	By comparing the genome sequences of cattle and peas, it has been determined that they have common genes.	.919	.129	.665	1.172	.302
Item 19	The DNA obtained as a result of combining DNA fragments from two different living things is called Recombinant DNA.	.762	.131	.504	1.019	.375
Item 1	By gene transfer it is possible for plants to produce animal proteins by gene transfer.	.304	.140	.029	.579	.125
Item 30	Enzymes are responsible for cutting and joining DNA.	.262	.141	-.015	.539	.376
Item 16	The question of knowing the information on the human genome and who has the right to examine it is a matter of biosecurity.	.220	.143	-.060	.499	.377
Item 36	The basic gene cloning workflow consists of determining the gene-isolation of DNA fragments to be cloned- insertion of isolated DNA and multiplication.	.133	.145	-.152	.417	.427
Item 35	Events such as mutation and crossing over are natural events that cause genetic changes without human intervention.	.088	.146	-.199	.374	.247
Item 28	Microorganisms obtained by genetic engineering can be used to clean toxic wastes in the environment.	.019	.148	-.272	.310	.338
Item 31	Foods obtained from genetically modified organisms can cause allergic reactions.	-.004	.149	-.297	.288	.439
Item 9	Stem cells are cells that can transform into many types of cells and have the ability to divide continuously.	-.028	.150	-.322	.266	.370
Item 33	Determination of paternity, determination of genetic diseases and similar processes can be done by DNA analysis.	-.028	.150	-.322	.266	.334
Item 18	One purpose of gene transfer to tomatoes is to extend their shelf life.	-.101	.152	-.400	.198	.456
Item 10	One of the aims of the biosafety law is to prevent the risks that may arise from organisms and their products obtained using modern biotechnology.	-.255	.158	-.565	.055	.410
Item 29	If modern biotechnological methods are not done in the right way, they can threaten the future of the world.	-.255	.158	-.565	.055	.405
Item 14	Organisms that have artificially altered one or more genes are called genetically modified organisms.	-.513	.170	-.845	-.180	.484
Item 2	Developing DNA technology may pose significant ethical problems.	-.544	.171	-.879	-.208	.276
Item 7	DNA technology methods allow us to identify genetic diseases even when the baby is in the womb.	-.977	.196	-1.362	-.592	.367

Table 4 shows the percentages of school-based correct responses for item 13, the most difficult item on the scale. It is seen that more students in HS3 answered Item 13 than in HS1 and HS2, respectively.

Table 4. School-based correct responses for Item 13

		HS1	HS2	HS3	Total	Item 13
Correct Answers to Item 13	N	116	41	124	281	By comparing the genome sequences of cattle and peas, it has been determined that they have common genes.
	%within Item 13	41.3%	14.6%	44.1%	100.0%	
	%within school	70.3%	53.2%	84.9%	72.4%	
	% of total	29.9%	10.6%	32.0%	72.4%	

Table 5 shows the percentages of school-based correct responses for item 7, the easiest item on the scale. It is seen that more students in HS3 answered Item 7 than in HS1 and HS2, respectively. The schools' ranking in terms of the percentage of correct responses to item 7 did not change, but the percentage gap between them decreased. In addition to these findings, the order of the percentage of correct answers in all the other items except for the Items 18, 28 and 29 is HS3 > HS1 > HS2, while the order of the percentage of correct answers for these three items is HS1 > HS3 > HS2.

Table 5. School-based correct responses for Item 7

		HS1	HS2	HS3	Total	Item 7
Correct Answers to Item 7	N	156	63	142	361	DNA technology methods allow us to identify genetic diseases even when the baby is in the womb.
	% within Item 7	43.2%	17.5%	39.3%	100.0%	
	% within school	94.5%	81.8%	97.3%	93.0%	
	% of total	40.2%	16.2%	36.6%	93.0%	

Comparisons of BBKS Scores by School, Grade and Gender

A factorial (three-way) ANOVA test was used to compare students' BBKS scores based on their school (HS1, HS2, HS3), grade (10th grade, 11th grade, 12th grade), and gender (male, female). The descriptive statistics findings indicated that students from HS3 (M= 15.29, SD= 2.37) gained higher scores on average than students from HS1 (M=14.67, SD= 2.65) and HS2

(M=11.68, SD= 2.94). Also, female students (M=14.77, SD=2.45) gained higher scores on average than male students (M=13.77, SD=3.18). Students in 12th grade (M=14.49, SD=2.76) gained higher scores on average than students in 11th grade (M=14.33, SD=2.83) and 10th grade (M=14.27, SD=2.89).

Three-Way ANOVA results indicated a significant main effect for the school effect ($F_{2,380}$= 50.91, p=.0001, η^2=.21). There were no statistically significant differences for grade ($F_{2,380}$=0.17, p.0.85) and gender ($F_{2,380}$=1.92, p=0.17) variables in terms of their biotechnology knowledge scores. Table 6 presents the Three-Way ANOVA analysis results for BBKS.

Table 6. Three-Way ANOVA Analysis Results for BBKS

Source	Type III Sum of Squares	Df	Sum of Squares	F	p-value
Corrected Model	801.356a	7	114.479	18.563	.0001
Intercept	36474.078	1	36474.078	5914.218	.0001
School	627.951	2	313.976	50.911	.0001
Grade	2.080	2	1.040	.169	.845
Gender	11.851	1	11.851	1.922	.166
Error	2343.530	380	6.167		
Total	82590.000	388			
Corrected Total	3144.887	387			

LSD post-hoc test results regarding the source of the observed difference in the school variable indicated that students from HS3 had significantly higher scores than students from HS1 (p=0.028) and HS2 (p=0.001). Likewise, students from HS1 had significantly higher scores than students from HS2 (p=0.001). Table 7 and Figure 1 present the LSD post-hoc results.

Table 7. LSD post-hoc test results for school variable

(I) school	(J) school	Mean Difference (I-J)	Std. Error	p-value
HS1	HS2	2.9784*	.34274	.001
	HS3	-.6210*	.28217	.028
HS2	HS1	-2.9784*	.34274	.001
	HS3	-3.5994*	.34976	.001
HS3	HS1	.6210*	.28217	.028
	HS2	3.5994*	.34976	.001

The main effects for the grade ($F_{2,380}$= 0.17, p=.845) and gender ($F_{1,380}$= 1.92, p=.17) were not statistically significant. However, the grade*gender interaction was statistically significant ($F_{2,380}$= 4.09, p=.018, η^2=.02). As can be seen in Figure 1, while females in 10th and 11th grade tend to obtain higher scores than males, males obtained higher scores than females in 12th grade.

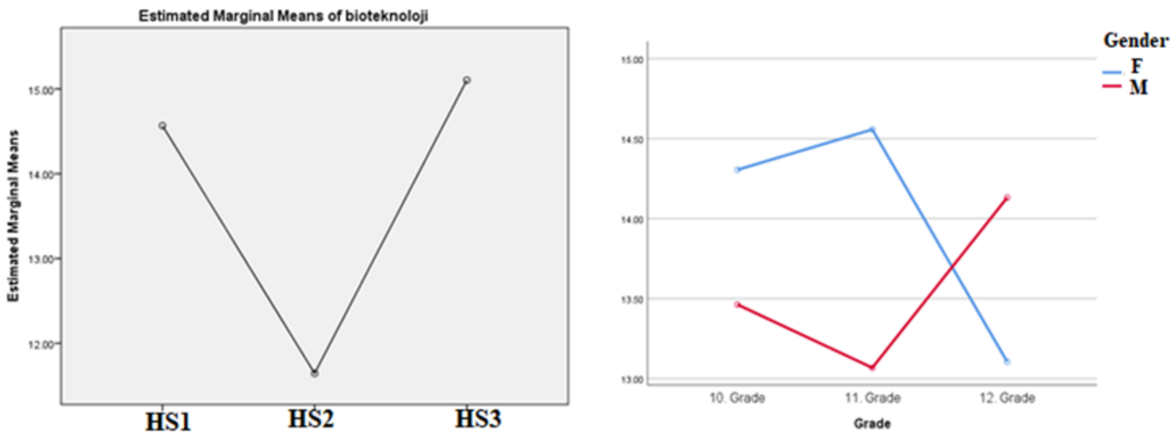


Figure 1. Estimated Marginal Means of BBKS by school and gender

Conclusions and Discussion

Validation of Basic Biotechnology Knowledge Scale (BBKS)

Being able to make use of the opportunities biotechnology offers depends largely on the accuracy and adequacy of the acquired knowledge. Evaluation of the accuracy and adequacy of the information requires the existence of accurate, valid and reliable measurement tools. In this study, a standardized scale in biotechnology and gene engineering (BBKS) was developed for all individuals who have completed secondary education. Rasch Measurement Model was used in the development of the scale and analysis of the data to compare gender, school and grade.

BBKS resulted in a 3-sub-topic, 17-item True-False Type Knowledge scale. According to Rasch analysis results, the most difficult item in the scale is "Item 13", and the easiest item is "Item 7". Also, "Item 31" is found as the closest to medium difficulty. Item 13 was the most difficult item on the scale, as students might not have an evolutionary perspective to understand that the ancestors of animals and plants are commonly based on their low apparent similarities. Item 7, on the other hand, was the easiest, as the information in the item can be frequently encountered in social media and daily life experiences. In some studies, the positive effects of social media on learning are also expressed (Özgen, Güngör, Emiroğlu, & Taş, 2007; Sıcaker, 2013).

When various instruments in many studies (Agaç, 2019; Arvanitayannis & Kystallis, 2005; Bayoğlu & Özgen, 2010; Bilen & Özel, 2012; Demir & Pala, 2007; Ergin, Gürsoy, Öcek & Çiçeklioğlu, 2008; Gürkan & Kahraman, 2018; Gürkan & Kahraman, 2019; Keskin, 2003; Keskin et al., 2010; Koçak, Türker, Kılıç, & Hasde, 2010; Konak & Hasancebi, 2021; Olsher & Dreyfus, 1999; Öcal, 2012; Özdemir, Güneş, & Demir, 2010; Özgen et al., 2007a, b; Priest et al., 2003; Prokop et al., 2007; Sönmez & Pektaş, 2017; Subrahmanyan & Cheng, 2000; Sürmeli & Şahin, 2009; Wie, Strohbehn, & Hsu,1998; Yılmaz & Öğretmen, 2014; Yüce & Yalçın, 2012) in the field of biotechnology education are examined, it is seen that BBKS differs from these studies in terms of the target audience, scope, type of data collection tool, and data analysis methods. For example, the Biotechnology Knowledge Test prepared by Yüce and Yalçın (2012) is different in terms of the target audience since they examined pre-service science teachers' biotechnology knowledge. Sönmez and Pektaş (2019) examined the effect of extracurricular activities on middle school students' views of the nature of science and biotechnology knowledge using Prokop et al.'s (2007) 16 Likert-type questions and also requested to explain their answers. In a different scale developed by Fonseca et al. (2012) to make a multidimensional analysis of secondary school students' perceptions of biotechnology, knowledge questions, mostly true/false questions, were also included. These knowledge questions contained items suitable for the topics of the developed scale, but there are differences in the distribution of items under the topics. This study is similar to our study in terms of the sample, but the scale also includes different dimensions apart from the knowledge test. In another study conducted in Slovenia (Paš, Vogrinc, Raspor, Udovč Knežević, & Čehovin Zajc, 2019), in the first stage, content analysis was conducted on 15 biotechnology topics selected from the entire high school curriculum, and in the second stage, a measurement tool was designed to determine students' knowledge of traditional and modern biotechnology. The measurement tool applied to high school students in the 17-18 age group was compiled from the questions in previous studies and consists of two parts. In the first part, knowledge of 18 modern biotechnology and seven traditional biotechnology items, and in the second part, attitudes towards modern biotechnology and biotechnology products were tried to be examined.

A scale also developed in Turkey, which is very similar to BBKS, was presented by Sıcaker et al. (2020). This study was prepared according to the secondary school biology curriculum. Curriculums have difficulty keeping up with the pace of development and change in biotechnology and may take some time to update. Some topics that need to be known today may take their place in the curricula over time. In this respect, there is a fundamental difference between BBKS and this scale; when there is a need for a scale to measure biotechnology knowledge at the secondary education level, one of the two scales can be preferred in line with the purposes of the studies to be conducted, and it provides an opportunity for the researchers in this respect. In addition, knowledge measured with a few questions in the previous study can be measured with a single item in the present study. In the present study, the subject has been handled in a more general structure since it was prepared in order to question the biotechnology knowledge of an individual who graduated from different secondary education departments (social field, science field, sports field, fine arts field, etc.). In addition, the fact that it consists of 17 items provides ease of answering. It is hoped that BBKS, which was developed without being completely dependent on the curriculum, will also lead to a development in the direction of making changes in the curriculum by noticing the deficiencies that can be seen with the items.

Comparison of High School Students' Biotechnology Knowledge with BBKS in terms of School, Gender, and Grade

Comparison of school, gender and grade results indicated that students attending the science-intensive high school (HS3) gained higher scores than HS1 and HS2, and the difference between their knowledge of biotechnology was significant. Considering the highest score that can be obtained from the scale is 17, it is obvious that students from HS3 and HS1 gained higher scores from BBKS; however, HS2 was lower than the other two schools. Here, it is possible to say that students' scores are directly proportional to their high school entrance scores since HS3 and HS2 require higher scores to enter. Also, it is possible to say that students' total average score is above average (M=13,88). These results are not consistent with other biotechnology knowledge studies. For example, Chen and Raffan (1999), in their study of 352 post-16 students studying in England and Taiwan, stated that the students had limited biotechnology knowledge in terms of the meaning and examples of genetic engineering. Similarly, Yüce and Yalçın (2012) showed that the biotechnology education pre-service science teachers received at high schools did not provide them with sufficient and permanent knowledge, while university education provided them with a medium level of knowledge. A recent study by de la Hoz et al. (2022) indicated that Swedish and Spanish pre-service primary school teachers showed a lack of knowledge about basic genetics that could negatively influence their ability to address biotechnological applications in their teaching. Since BBKS is aimed at examining students' basic knowledge, the higher scores they gained did not come as surprising and showed that BBKS is an appropriate scale to examine high school graduates' biotechnology knowledge regardless of gender, grade, and school.

In terms of the items in BBKS, for 14 out of 17 items, the answering rate of schools resulted as HS3>HS2 and HS1; however, for items 18, 28 and 29, the answering rate was HS2>HS3> HS1. This result might be the consequence of having 12th-grade students. This might be the result of two situations. Firstly High School Biology curriculum in Turkey adopts a spiral curriculum approach, and students get more detailed knowledge of biotechnology as they pass to 12th grade. Secondly, the students in 12th grade are preparing for university entrance exams, which require them to review their previous lessons. The main effect of gender was not statistically significant. However, the grade*gender interaction was statistically significant, favouring males with a small effect size. This observed effect was possibly due to the uneven sample size of 12th-grade students. More studies with 12th-grade samples are needed to reveal whether the grade*gender interaction observed in the current study exists in the population of 12th-grade Turkish High School students.

With the developed Basic Biotechnology Knowledge Scale (BBKS), it will be possible to determine the knowledge level of individuals and, accordingly, their deficiencies related to the subject. In this way, it is thought that it can be a guide for the improvement of high school programs. This scale will contribute to the achievement of distant goals in biotechnology education.

Limitations of the Study-Future Research

Although there were questions about vaccines and microorganisms in the early stages of scale development, the absence of questions on this subject in the final form of the scale was regarded as a limitation of BBKS. It is important to re-evaluate the questions about vaccines and viruses after the SARS-CoV-2 pandemic to eliminate the limitations of the scale.

Furthermore, this study is limited to the Turkish national setting. As a result, future research should broaden the scope of the study to evaluate the generalizability of the findings in various educational and cultural situations.

Temel Biyoteknoloji Bilgi Ölçeğinin (TBBÖ) Rasch Ölçüm Modeline Göre Geliştirilmesi

Özet:

Bu çalışmanın amacı, temel biyoteknoloji bilgisini ölçmek için bir ölçek geliştirmek, ölçeğin psikometrik özelliklerini incelemek ve öğrencilerin test performanslarında okul, sınıf ve cinsiyete göre farklılık olup olmadığını incelemektir. Bu ölçeğin geliştirme aşaması Türkiye'nin batısındaki bir ilde 388 lise öğrencisi örneklemi ile gerçekleştirilmiştir. Ölçeğin psikometrik özellikleri Rasch modeli kullanılarak incelenmiştir. 17 maddeden oluşan son ölçeğin K-R iç tutarlılık katsayısı 0,77 olarak hesaplanmıştır. Madde-toplam korelasyonlarının bir madde (Madde 1, 0.13) dışında 0.25 ile 0.48 arasında değiştiği görülmüştür. Rasch analizinin sonuçları, ölçeğin Rasch modeline uyduğunu ve düşük ve yüksek performanslı sınava girenleri ayırt edebildiğini göstermiştir. Üç Yönlü ANOVA sonuçları, okul değişkeni için önemli bir ana etki göstermiştir Ayrıca, biyoteknoloji bilgi puanları açısından sınıf ve cinsiyet değişkenleri arasında istatistiksel olarak anlamlı bir fark bulunamamıştır. Bununla birlikte, sınıf*cinsiyet etkileşimi, küçük etki boyutuna sahip erkeklerin lehine istatistiksel olarak anlamlı bulunmuştur. Gözlenen bu etkinin, 12. sınıf öğrencilerinin eşit olmayan örneklem büyüklüğünden kaynaklandığı düşünülmektedir. Bu da Temel Biyoteknoloji Bilgisi Ölçeğinin (TBBÖ) lise öğrencilerinin bilgi düzeylerini değerlendirmek için kullanılabileceğini göstermektedir

Anahtar kelimeler: biyoteknoloji, bilgi ölçeği, Rasch ölçüm modeli, ölçek geliştirme

References

Acarlı, D. S. (2016). Determining prospective biology teachers' cognitive structure in terms of "biotechnology". *Journal of Baltic Science Education, 15*, 494–505.

Ağaç, H. (2019). *Fen bilgisi öğretmen adaylarının "Tarımsal biyoteknoloji" konusundaki yapılandırılmış deney uygulamalarının bilgi ve tutumlarına etkisi* [The effect of science teacher candidates on knowledge and attitudes of structured experimental applications on "agricultural biotechnology"]. (Unpublished master's thesis). Canakkale Onsekiz Mart University Institute of Educational Sciences, Canakkale.

Akman, S. B. (2007). *Avrupa Birliği'nin biyoteknolojik ürün ve uygulamalara yönelik tüketici politikası ve Türkiye'nin uyumu* [Consumer policy towards biotechnological applications and products in the European Union and harmonisation of Turkey]. (Unpublished master's thesis). Ankara University Insitute of Biotechnology, Ankara.

Altıntaş, Ö., & Kutlu, Ö. (2019). Investigating differential item functioning of Ankara University examination for foreign students by recursive partitioning analysis in the Rasch model. *International Journal of Assessment Tools in Education*, 6(4), 602–616. https://doi.org/10.21449/ijate.554212

American Educational Research Association, American Psychological Association, National Council on Measurement in Education, Joint Committee on Standards for Educational & Psychological Testing (US). (2014). *Standards for educational and psychological testing.*

Washington, DC: American Educational Research Association. Retrieved from https://www.apa.org/science/programs/testing/standards

Arık, R. S., & Kutlu, Ö. (2013). Scaling primary school teachers' competence based on judgmental decisions in the field of measurement and evaluation. *Journal of Educational Sciences Research*, 3(2), 163-196. https://dergipark.org.tr/tr/pub/ebader/issue/44712/555605

Atasoy, B., Atıcı, T., Taşar, M. F. & Taflı, T. (2020). Development and validation of biotechnology knowledge scale (BKS). *Hellenic Journal of STEM Education*, 1(1), 33-42. doi: 10.51724/hjstemed.v1i1.3

Ayar, A., & Hasipek, S. (2003). Genetik modifiye gıdalar [Genetically Modified Foods]. *Standard: Technical and Economic Journal*, 42(494),73-79.

Bal, Ş., & Keskin, N. (2002, September). *Grup tartışması yoluyla öğrencilerin genetik mühendisliği uygulamaları ile ilgili tutum ve görüşlerinin değerlendirilmesi [Evaluation of students' attitudes and opinions about genetic engineering applications through group discussion]*. Paper presented at the 5th National Congress of Science and Mathematics Education, Ankara.

Baştürk, R. (2010). Bilimsel araştırma ödevlerinin çok yüzeyli Rasch ölçme modeli ile değerlendirilmesi [Evaluation of research assignments with many-facet Rasch model]. *Journal of Measurement and Evaluation in Education and Psychology*, 1(1), 51-57.

Bayoğlu, A. S., & Özgen, Ö. (2010). An examination about the consumer attitudes towards biotechnology and perceived benefit and perceived risk. *The Journal of International Social Research*, 3(10), 90-103.

Bilen, K., & Özel, M. (2012). Üstün yetenekli öğrencilerin biyoteknolojiye yönelik bilgileri ve tutumları [Gifted students' knowledge of and attitudes toward biotechnology]. *Necatibey Faculty of Education Electronic Journal of Science and Mathematics Education*, 6(2), 135-152.

Bonfadelli, H. (2005). Mass media and biotechnology: knowledge gaps within and between European countries. *International Journal of Public Opinion Research*, *17*(1), 42–62. doi:10.1093/ijpor/edh056

Boone, W. J. (2016). Rasch analysis for instrument development: why, when, and how? *CBE Life Science Education*, 15: rm4, 1-7. doi: 10.1187/cbe.16-04-0148.

Boone, W. J., & Scantlebury, K. (2006). The role of Rasch analysis when conducting science education research utilizing multiple-choice tests. *Science Education*, 90(2), 253–269. doi:10.1002/sce.20106

Borgerding, L. A., Sadler, T. D., & Koroly, M. J. (2013). Teachers'concerns about biotechnology education. *Journal of Science Education and Technology,* 22, 133–147.

Casanoves,M., González,A., Salvadó, A., Haro, J., & Novo, M. (2015). Knowledge and attitudes towards biotechnology of elementary education pre-service teachers: the first Spanish experience. *International Journal of Science Education,* 37(17), 2923-2941, doi: 10.1080/09500693.2015.1116718

Chabalengula, V. M., Mumba, F., & Chitiyo, J. (2011). American elementary education pre-service teachers' attitudes towards biotechnology processes. *International Journal of Environmental and Science Education,* 6(4), 341-357.

Chen, S. Y., & Raffan, J. (1999). Biotechnology: student's knowledge and attitudes in the UK and Taiwan. *Journal of Biological Education,* 34(1), 17–23.

Cohen, L., Manion, L., & Morrison, K. (2018). *Research methods in education* (8th Edition). NY: Routledge.

Dawson, V. (2007). An exploration of high school (12–17-year-old) students' understandings and attitudes towards biotechnology processes. *Research in Science Education,* 37, 59-73.

Dawson, W., & Schibeci, R. (2003). Western Australian school students' understanding of biotechnology. *International Journal of Science Education,* 25(1), 57–69.

Demir, A., & Pala, A. (2007). Genetiği değiştirilmiş organizmalara toplumun bakış açısı [Perceptions of society towards genetically modified organisms]. *Hayvansal Üretim,* 48(1), 33-43.

de la Hoz, M.C., Solé-Llussà, A., Haro, J., Gericke, N. & Valls, N. (2022). Student primary teachers' knowledge and attitudes towards biotechnology—are they prepared to teach biotechnological literacy?. *Journal of Science Education and Technology, 31,* 203–216 (2022). https://doi.org/10.1007/s10956-021-09942-z

Elhan, A. H. & Atakurt, Y. (2005). Ölçeklerin değerlendirilmesinde niçin Rasch analizi kullanılmalıdır? . *Ankara Üniversitesi Tıp Fakültesi Mecmuası* , 58 (1) , 47-50 . DOI: 10.1501/Tipfak_0000000134

Ergin, I., Gürsoy, Ş. T., Öcek, Z. A., & Çiçeklioğlu, M. (2008). Knowledge, attitude and behavior of medical technology vocational training school students about genetically modified organisms. *TAF Preventive Medicine Bulletin,* 7, 503-508.

Fonseca, M. J., Costa, P., Lencastre, L., & Tavares, F. (2012). Disclosing biology teachers' beliefs about biotechnology and biotechnology education. *Teaching and Teacher Education,* 28, 368–381. doi: 10.1016/j.tate.2011.11.007

Frewer, L. J., Howard, C., & Shepherd R. (1997) Public concerns in the United Kingdom about general and specific applications of genetic engineering: risk, benefit, and ethics. *Science Technology Human Values,* 22 (1), 98-124. doi: 10.1177/016224399702200105

Frewer, L. J., Shepherd, R., & Sparks, P. (1994) The interrelationship between perceived knowledge, control and risk associated with a range of food-related hazards targeted at the individual, other people and society. *Journal of Food Safety,* 14 (1), 19-40. doi: 10.1111/j.1745-4565.1994.tb00581.x

Gardner, G. E., & Jones, M. G. (2011). Science instructors' perceptions of the risks of biotechnology: implications for science education. *Research in Science Education,* 41(5), 711-738. doi:10.1007/s11165-010-9187-0

Gardner, G., Jones, G., Taylor, A., Forrester, J., & Robertson, L. (2010). Students' risk perceptions of nanotechnology applications: implications for science education. *International Journal of Science Education,* 32(14), 1951-1969. doi:10.1080/09500690903331035

Gaskell, G., Stares, S., Allansdottir, A., Allum, N., Corchero, C., & Jackson, J. (2006) *Europeans and biotechnology in 2005: Patterns and trends: Final report on Eurobarometer 64.3*. Publications Office of the European Union, Luxembourg.

Gelamdin, R. B., Alias, N., & Attaran, M. (2013). Students' and teachers' perspectives on biotechnology education: A review on publications in selected journals. *Life Science Journal,* 10(1), 1210–1221.

Gonzalez, A., Casanoves, M., Salvado, Z., Barnett, J., & Novo, M. T. (2013). Biotechnology literacy: much more than a gene story. *The International Journal of Science in Society,* 4, 27–35.

Gunter, B., Kinderlerer, J., & Beyleveld, D. (1998). Teenagers and biotechnology: a survey of understanding and opinion in Britain. *Studies in Science Education,* 32, 81–112.

Gürkan, G., & Kahraman, S. (2018). Fen bilgisi öğretmen adaylarının biyoteknoloji ve genetik mühendisliği bilgi düzeyleri [Knowledge levels of preservice science teachers about biotechnology and genetic engineering]. *Karaelmas Eğitim Bilimleri Dergisi, 6*(1), 25-39.

Gürkan, G. & Kahraman, S. (2019). Fen bilgisi öğretmenlerinin biyoteknoloji ve genetik mühendisliği bilgi düzeylerinin çeşitli değişkenler açısından incelenmesi: Malatya ili örneği [Investigation of knowledge levels of science teachers on biotechnology and genetic engineering in terms of different variables: The case of Malatya province]. *Inonu University Journal of the Graduate School of Education*, 6(12), 66-78.

Hanegan, N. L., & Bigler, A. (2009) Infusing authentic inquiry into biotechnology. *Journal of Science Education and Technology*, 18, 393-401.

Harms, U. (2002). Biotechnology education in schools. *Electronic Journal of Biotechnology, 5*(3). Retrieved June 21, 2021, from http://ejbiotechnology.info/index.php/ejbiotechnology/article/view/v5n3-i03/980

Hebebci, M. T., & Shelley, M. (2018). Analysis of the relationship between university students' problematic internet use and loneliness. *International Journal of Assessment Tools in Education*, *5*(2), 223-234. Doi: 10.21449/ijate.402690

Ho, M. W. (2001). *Genetik mühendisliği: Rüya mı kabus mu*? [Genetic engineering: Dream or nightmare?]. İstanbul: Türkiye İş Bankası Kültür Yayınları.

Kahveci, D., & Özçelik, B. (2008). Attitudes of Turkish consumers towards genetically modified foods. *International Journal of Natural and Engineering Sciences*, 2(2), 53-57.

Kaptan, F. (1994). Rasch modeli madde parametrelerini kullanarak en yüksek olabilirlik yöntemiyle yetenegin kestirilmesi [Estimation of ability with the highest likelihood method using Rasch model item parameters]. *Hacettepe University Journal of Education,* 10, 95-97.

Kaskatı, O. T. (2011). *Rasch modelleri kullanarak romatoid artirit hastaları özürlülük değerlendirimi için bilgisayar uyarlamalı test yönteminin geliştirilmesi* [Development of computer adaptive testing method using with Rasch models for assessment of disability in rheumatoid arthritis patients]. (Unpublished doctoral dissertation). Ankara University Institute of Health Sciences, Ankara.

Keskin, N. (2003*). Fen bilgisi eğitimi 3. sınıf öğrencilerinin gen klonlama konusunu öğrenmelerine poster sunumu etkinliğinin etkisi* [The effect of poster presentation activity on science education 3rd grade students' learning gene cloning]. (Unpublished master’s thesis*)*, Gazi University Institute of Educational Sciences, Ankara.

Keskin, Y., Lüleci, N. E., Özyaral, O., Altıntaş, Ö., Sağlık, A., Lisar, H., Turan, A., & Top, Y. (2010). Maltepe üniversitesi tıp fakültesi örencilerinin genetiği değiştirilmiş organizmalar hakkında bilgi tutum ve davranışları [Knowledge, attitude and behavior of the students in medical school of Maltepe University about genetically modified organisms]. *Maltepe Medical Journal*, 2 (1), 14-23.

Klop, T., & Severiens, S. (2007). An exploration of attitudes towards modern biotechnology: A study among Dutch secondary school students. *International Journal of Science Education,* 29(5), 663-679.

Koçak, N., Türker, T., Kılıç, S., & Hasde, M. (2010). Assessment of knowledge, attitude, and behavior level of medical school students about genetically modified organisms. *Gulhane Medical Journal*, 52(3), 198-204.

Konak, M. A., & Hasancebi, S. (2021). Evaluation of biology teachers' knowledge and attitudes towards biotechnology and its applications. *Journal of Instructional Technologies and Teacher Education,* 10 (1), 1-15. doi: 10.51960/jitte.826174.

Koparan, T., & Güvenen, B. (2013). The effect of project-based learning approach on primary school students' statistical literacy levels about sample concept. *Journal of Research in Education and Teaching,* 2(1),185-196.

Kustiana, Suratno & Wahyuni, D. (2020). The analysis of metacognitive skills and creative thinking skills in STEM education at senior high school for biotechnology. *Journal of Physics: Conference Series*,1465,012045. doi:10.1088/1742-6596/1465/1/012045

Lamanauskas, V., & Makarskaitė-Petkevičienė, R. (2008). Lithuanian university students' knowledge of biotechnology and their attitudes to the taught subject. *Eurasia Journal of Mathematics, Science and Technology Education,* 4(3), 269-277.

Lock, R., & Miles, C. (1993). Biotechnology and genetic engineering: Students' knowledge and attitudes. *Journal of Biological Education,* 27(4), 267–273.

Lyson, T. A. (2002). Advanced agricultural biotechnologies and sustainable agriculture. *Trends in Biotechnology*, 20, 193-196.

Massarani, L., & Moreira, I. C. (2005). Attitudes towards genetics: A case study among Brazilian high school students. *Public Understanding of Science*, 14, 201-212. doi:10.1177/0963662505050992

Ministry of National Education (MoNE) (2018a). Fen bilimleri dersi öğretim programı (İlkokul ve ortaokul 1,2,3,4,5,6,7 ve 8. sınıflar) [Science course curriculum (primary and secondary school 1,2,3,4,5,6,7 and 8th grades]. Ankara. Retrieved from http://mufredat.meb.gov.tr/Programlar.aspx.

Ministry of National Education (MoNE) (2018b). Biyoloji Dersi Öğretim Programı [Biology Curriculum]. Ankara. Retrieved from http://mufredat.meb.gov.tr/Dosyalar/20182215535566Biyoloji%20d%C3%B6p.pdf

Morris, S. H., & Adley, C.C. (2001). Irish public perceptions and attitudes to modern biotechnology: an overview with a focus on GM foods. *Trends in Biotechnology*, 19(2), 43-48.

Mowen, D. L., Roberts, T. G., Wingenbach, G. J., & Harlin, J. F. (2006). Barriers, roles, and information source preferences of agricultural science teachers' for teaching biotechnology: A case study. *Proceedings of the 33rd Annual National Agricultural Education Research Conference*, Charlotte, NC.

Nguyen, T. H., Han, H. R., Kim, M. T., & Chan, K. S. (2014). An introduction to item response theory for patient-reported outcome measurement. *The patient*, 7(1), 23–35. doi:10.1007/s40271-013-0041-0

Olsher, G., & Dreyfus, A. (1999). The "ostension-teaching" approach as a means to develop junior-high student attitudes towards biotechnologies. *Journal of Biological Education*, 34(1), 25-31.

Öcal, E. (2012). *İlköğretim fen bilgisi öğretmenlerinin biyoteknoloji (genetik mühendisliği) farkındalık düzeyleri* [The level of biotechnology (genetic engineering) awareness of elementary science teachers]. (Unpublished master’s thesis). Inonu University Institute of Educational Sciences, Malatya.

Ömür, S., & Erkuş, A. (2013). Dereceli puanlama anahtarıyla, genel izlenimle ve ikili karşılaştırmalar yöntemiyle yapılan değerlendirmelerin karşılaştırılması [Comparison of the evaluations which were done with rubric, overall impression and paired comparisons]. *Hacettepe University Journal of Education, 28*(2), 308-320.

Özdemir, O., Güneş, M. H., & Demir, S. (2010). The level of knowledge-attitudes of students towards genetically modified organism (GMO) and its evaluation according to sustainable consumption education. *Ondokuz Mayis University Journal of Education Faculty*, 29 (1), 53- 68.

Özgen, Ö. (1995). Biyoteknoloji ve tüketici [Biotechnology and consumer]. *Verimlilik Dergisi [Journal of Productivity]*, 2, 141-147.

Özgen, Ö., Emiroğlu, H., Yıldız, M., Taş, A. S., & Purutçuoğlu, E. (2007a). *Tüketiciler ve modern biyoteknoloji: Model yaklaşımlar* [Consumers and modern biotechnology: Model approaches]. Ankara Üniversitesi Biyoteknoloji Enstitüsü Yayınları [Ankara University Institute of Biotechnology Publications], Ankara Üniversitesi Basımevi, Ankara.

Özgen, Ö., Güngör, N., Emiroğlu, H., & Taş, A. S. (2007b, May). *College students' opinions about consumer education and information sources towards biotechnological* *applications and products.* Paper presented at the 5th International Conference on Communication and Mass Media, Athens, Greece.

Pas, M., Vogrinc, J., Raspor, P., Knezevic, N. U., & Zajc, J. C. (2019). Biotechnology learning in Slovenian upper-secondary education: Gaining knowledge and forming attitudes. *Research in Science & Technological Education*, 37(1), 110-125. doi: 10.108 0/02635143.2018.1491473.

Priest, S. H., Bonfadelli, H., & Rusanen, M. (2003). The "trust gap" hypothesis: Predicting support for biotechnology across national cultures as a function of trust in actors. *Risk Analysis*, 23(4), 751–766.

Prokop, P., Leskova, A., Kubiatko, M., & Diran, C. (2007). Slovakian students' knowledge of and attitudes toward biotechnology. *International Journal of Science Education,* 29(7), 895-907.

Saefi, M., Fauzi, A., Kristiana, E., Adi, W. C., Muchson, M., Setiawan, M. E., Islami, N. N., Ningrum, D., Ikhsan, M. A., & Ramadhani, M. (2020). Survey data of covid-19-related knowledge, attitude, and practices among Indonesian undergraduate students. *Data in brief*, 31, 105855. doi:10.1016/j.dib.2020.105855

Schilling, B. J., Hallman, W., Hossain, F., & Adelaja, A. O. (2003). Consumer perception of food biotechnology: evidence from a survey of US consumers. *Journal of Food Distribution Research*, 34(1), 30–35.

Semerci, Ç. (2011a). Analyzing microteaching applications with many-facet Rasch measurement model. *Education and Science*, 36 (161), 14-25.

Semerci, Ç. (2011b). Doktora yeterlikler çerçevesinde öğretim üyesi, akran ve öz değerlendirmelerin Rasch ölçme modeliyle analizi [Analysis of faculty members, peers and self-assessments within the framework of doctoral qualifications with the Rasch measurement model]. *Journal of Measurement and Evaluation in Education and Psychology*, 2(2), 164-17.

Shaw, A. (2002). "It just goes against the grain." public understandings of genetically modified (GM) food in the UK. *Public Understanding of Science*, 11(3), 273-291.

Sıcaker, A. (2013). *The development study of a secondary school biotechnology and genetic engineering knowledge scale with Rasch measurement model* [The development study of a secondary school biotechnology and genetic engineering knowledge scale with Rasch measurement model]. (Unpublished master's thesis). Balıkesir University Institute of Sciences, Balikesir.

Sıcaker, A., & Öz Aydın, S. (2015). Ortaöğretim biyoteknoloji ve gen mühendisliği kavramlarının öğrenciler tarafından değerlendirilmesi [Evaluation of biotechnology and genetic engineering concepts by students]. *Ondokuz Mayis University Journal of Faculty of Education,* 34 (2)*,* 51- 67.

Sıcaker, A., Öz Aydın, S., & Saçkes, M. (2020). Ortaöğretim Biyoteknoloji Bilgi Testi (OBBT) [The secondary school biotechnology knowledge test (SBKT)]. *Pamukkale Üniversitesi Eğitim Fakültesi Dergisi, 50*, 426-448. doi:10.9779/pauefd.595078

Sjöberg, L. (1996*). Risk perceptions by politicians and the public*. Rhizikon: Risk Research Reports No. 23. Stockholm, Sweden: Center for Risk Research.

Sönmez, E., & Pektaş, M. (2017). Ortaokul öğrencilerine müfredat dışında uygulanan bazı biyoteknoloji etkinliklerinin bilimin doğası görüşleri ve biyoteknoloji bilgilerine etkisi [The effects of some activities of biotechnology in extra-curricular on middle school students' nature of science perceptions and biotechnology knowledge]. *Kastamonu Education Journal*, 25(5), 2019-2036.

Sparks, P., & Shepherd, R. (1994). Public perceptions of the potential hazards associated with food production and food consumption: An empirical study. *Risk Analysis,* 14(5), 799-806. doi:10.1111/j.1539-6924.1994.tb00291.x

Subrahmanyan, S., & Cheng, P. S. (2000). Perceptions and attitudes of Singaporeans toward genetically modified foods. *Journal of Consumer Affairs,* 34(2), 269-290.

Sürmeli, H., & Şahin, F. (2009). Üniversite öğrencilerinin biyoteknoloji çalışmalarına yönelik bilgi ve görüşleri [Knowledge and opinions of university students on biotechnology studies]. *Cukurova University Faculty of Education Journal,* 37*,* 33-45.

Sürmeli, H., & Şahin, F. (2010a). University students' attitudes towards biotechnological studies. *Education and Science*, 35(155), 145-157.

Sürmeli, H., & Şahin, F. (2010b). University students' bioethical perceptions about genetic engineering: genetic testing and genetic diagnosis. *Journal of Turkish Science Education,* 7(2), 119-132.

Testa, I., Costanzo, G., Crispino, M., Galano,S., Parlati,A., Tarallo, O. Tricò, F., & di Uccio, U.S. (2022). Development and validation of an instrument to measure students' engagement and participation in science activities through factor analysis and Rasch analysis. *International Journal of Science Education,* 44(1), 18-47. doi: 10.1080/09500693.2021.2010286

Turan, M., & Koç, I. (2012). Pre-service science teachers' attitudes towards biotechnology applications, *Trakya University Journal of Education,* 2(2), 74-83.

Tyas, A., Senam, Wiyarsi, A., & Laksono, E. (2020). Chemistry teaching self-efficacy: A scale development. *Journal of Physics: Conference Series, 1440*, 012005, 1-8. doi:10.1088/1742-6596/1440/1/012005

Wie, S. H., Strohbehn, C. H., & Hsu, C. H. C. (1998). Iowa dietitians' attitudes toward and knowledge of genetically engineered and irradiated foods. *Journal of the American Dietetic Association*, 98(11), 1331–1333.

Willis, G. B. (2005). *Cognitive interviewing: A tool for improving questionnaire design: A tool for improving questionnaire design*. Thousand Oaks, CA: Sage.

Woudstra, A. J., Smets, E. M., Galenkamp, H., & Fransen, M. P. (2019). Validation of health literacy domains for informed decision making about colorectal cancer screening using classical test theory and item response theory. *Patient Education and Counseling,* 102(12), 2335–2343. doi: 10.1016/j.pec.2019.09.016

Wright, B. D., & Mok, M. C. M. (2004). An overview of the family of rasch measurement models. Introduction to Rasch measurement. In E.V. Smith Jr. & R.M. Smith (Eds)., *Introduction to Rasch Measurement* (pp.1-24). Maple Grove, MN: JAM Press.

Wright, B. D., & Stone, M. H. (1979). *Best test design*. Chicago: Mesa Press.

Yılmaz, M., & Öğretmen, T. (2014). Biyoloji öğretmen adaylarının gen teknolojisine ilişkin bilgi düzeyleri ve bilgi kaynaklarının incelenmesi [The level of knowledge of prospective biology teachers on gene technology and their source of information]. *Pegem Journal of Education and Instruction*, 4(4), 59-76.

Yüce, Z., & Yalçın, N. (2012, June). *Fen bilgisi öğretmen adaylarının biyoteknoloji konusundaki bilgi düzeyleri* [Preservice science teachers' knowledge levels of biotechnology]. Paper presented at the 10th National Congress of Science and Mathematics Education, Niğde.

Necatibey Eğitim Fakültesi Elektronik Fen ve Matematik Eğitimi Dergisi
Cilt 16, Sayı 2, Aralık 2022, sayfa 381-405. ISSN: 1307-6086
Necatibey Faculty of Education Electronic Journal of Science and Mathematics Education
Vol. 16, Issue 2, December 2022, pp. 381-405. ISSN: 1307-6086

Research Article

Prospective Science Teachers' Arguments Regarding a Discrepant Event and Their Thoughts on Using Them in Science Education

Oktay KIZKAPAN

[1] Nevşehir Hacı Bektaş Veli University Department of Educational Sciences, Nevşehir/Turkiye, okizkapan@nevsehir.edu.tr, http://orcid.org/0000-0001-6781-9879

Received : 3.4.2022 Accepted : 5.11.2022

Doi:

Abstract – The aim of this study is to determine the prospective science teachers' (PSTs) arguments about a discrepant event and their views on the use of such discrepant events in science education. In the study, an instrumental case study, one of the qualitative research designs, was adopted. The study group of the research consisted of 73 prospective teachers studying at science education department of a stated university located in Central Anatolia Region of Turkey. The data of the research was collected by using written documents. The collected data in the study were analyzed using descriptive analysis. Analysis showed that the prospective science teachers' individual arguments about the discrepant event were weak and insufficient to explain the discrepancy. When prospective science teachers are provided with additional information and encouraged to cooperate effectively and allowed to work as a group, the quality and explanatory power of the arguments increased. However, it was observed that prospective teachers were not able to produce high-level arguments when they worked neither individually nor as a group.

Key words: discrepant events, science education, argumentation, chemistry education.

Introduction

Advent of Covid-19 pandemic changed educational environments dramatically. As a result of this change, distance education has become the reality of the lives of students and teachers. This reality has brought some problems together. One of these problems is lack of student interaction (Taşçı, 2021). Therefore, distance education environments should be student-centered. As a matter of fact, research show that teacher-centered teaching methods such as direct instruction and question-answer are mostly used in distance education

(Bakioğlu & Çevik, 2020). However, this situation must be changed. Therefore, there is a need for contents and activities that can be used in distance education environments, in accordance with the constructivist approach, and which will improve students' understanding by making them undergo a conceptual change. This need can be met if in-service and prospective teachers can carry out activities that will bring about conceptual change in students in educational environments.

According to Piaget, one of the advocates of cognitive theory, the cognitive equilibrium state of the individual deteriorates as a result of a new phenomenon. Then, with the explanation about the phenomenon that caused the deterioration, a more developed rebalance can be established (Senemoğlu, 2010). Individuals continue their learning in a cycle of balance, imbalance, and rebalance as a result of social and environmental interactions. Therefore, the main task of educational environments is to confront students with situations that will create a cognitive conflict.

According to the conceptual change theory which is a frequently discussed topic in science education (Harteis et al., 2020), the students need to stay in a cognitively contradictory condition for the conceptual change to take place. Science educators aim to teach the concepts that are accepted as scientifically correct and to change the conceptual structures that exist in students (Vaughn et al., 2020). Different models have been proposed for conceptual change. For example, while Sinatra et al. (2014) suggested adding emotional variables such as engagement, motivation and self-threat to conceptual change models, Nadelson et al. (2018) suggested a more contemporary dynamic model. According to conceptual change theory suggested by Posner et al. (1982), students should be exposed to phenomenon that contradict their ideas and should be aware of the inadequacies of their own explanations. Therefore, they should be exposed to cognitive conflicts (Appleton, 1996). In this way, students can search for alternative explanations to solve the contradictory situation caused by a natural effort to get rid of the contradictory situation they are in. Thus, they will learn as a result of cognitive conflicts and resolving these conflicts (Limón, 2001), which will motivate them to learn.

In the literature, there are different teaching methods, techniques and tools developed to confront students with situations that create a cognitive contradiction. One of these tools is discrepant events (González-Espada et al., 2010). Discrepant events defined as situations that are inconsistent with intuitive expectations (Mason et al., 2004) are effective tools that can be used to increase students' interest in science and their motivation to learn scientific principles and concepts more conceptually (Wright & Govindarajan, 1992). Discrepant events are used in physics, chemistry, biology, social sciences, and teacher training (O'Brien et al., 1994). Besides in these areas, "discrepant events are very powerful ways to stimulate interest, motivate students to challenge their covert science misconceptions, and promote higher-order thinking skills" (González-Espada et al., 2010: 508).

According to Suprapto (2020), there are five types of misconceptions as preconceived notions, non-scientific beliefs, conceptual misunderstandings, misconceptions of local languages and factual misconceptions. Besides; students, teachers, teaching materials or literature, context and teaching methods are the main reasons of these misconceptions. Therefore, to overcome misconceptions, the types of misconceptions and their resources should be considered and well-known by educators. As the types and sources of misconceptions are known, there are also solutions for how to eliminate the misconceptions. At this point, using discrepant events can be helpful because when students face a discrepancy, they show a strong desire to resolve it. By this desire, students start to search for possible explanations (Kavogli, 1992). When they find a satisfying explanation, they make a conceptual transformation in their minds. In this way, students do not have misconceptions while learning the new subjects, and they can also overcome their existing misconceptions. Also, when students realize a discrepancy in a scientific phenomenon, they start to question and re-think about the phenomenon. Accordingly, they need additional information and start to search to revise their explanations. Through these struggles, they stay focused on the concept they are trying to understand. As a result, students develop a more conceptual understanding of the content knowledge they aim to learn (Blikstein et al., 2016; Hewson & Hewson, 1984). Hence, it can be inferred that discrepant events are effective tools that can be used to develop students' content knowledge, avoid and eliminate misconceptions by creating conceptual contradictions.

Although it is known that exposing students to situations that create cognitive conflict is effective in science education, teachers face various issues in taking such activities to the classroom. Being worried not to be on time on the schedule, difficulty to control the classroom management during group activities, lack of experimental materials, and the tendency of distinguishing lessons as theoretical and experimental are some of these issues (de Oliveira & Fischer, 2017). Despite the teachers' concerns, discrepant events can be integrated into the argumentation processes effectively using the predict-observe-explain (POE) technique (e.g. Shemwell & Furtak, 2010). In using the POE technique, students are required to predict the outcomes of an experiment or demonstration, justify their prediction, observe the demonstration, and then clarify discrepancies between their prediction and observation. Studies have shown that POE is effective in developing students' skills (Sarah et al., 2021), increases students' problem-solving skills and self-efficacy (Fitriani et al., 2020) and students' POE attitudes positively predict students' self-confidence and critical attitudes (Hong et al., 2021). It is also known that the POE technique is widely used as an assessment tool for science education at different levels (White & Gunstone, 1992, cited in Karamustafaoğlu & Mamlok-Naman, 2015). Considering these aspects, the POE can be used to determine PSTs' arguments in discrepant event activities.

In online and traditional science learning environment, several approaches have been used to increase students' argumentation abilities. As one of these approaches, discrepant events can be used as effective tools in either online or face to face argumentative instructions, since they promote and support argumentation in science lessons. Discrepant events are productive contexts in which students are required to engage in argumentation to make sense of a given situation. Sampson and Clark (2009) state that when students craft convincing and persuasive arguments with the available data regarding a discrepant event, they learn about not only the content knowledge related to the discrepant event, but also about the argumentation. Therefore, discrepant events are used as a tool to help students to learn how to engage in scientific argumentation (Sampson & Clark, 2009; Sampson et al., 2011). Also, discrepant events and argumentation have a mutual effect on conceptual change. Both argumentation and discrepant events facilitate conceptual change in scientific knowledge. In other words, if teachers have adequate pedagogical knowledge about discrepant events and argumentation to utilize them in their science lessons, they can successfully bring about significant change in their students' perception towards science content knowledge (Anderson & Smith, 1983).

Discrepant events can be performed in the form of demonstration experiments as individual or small group activities. In this way, students' attention can be drawn, and they can be canalized to scientific research without requiring a long time (Mancuso, 2010). Therefore, PSTs need to be familiar with preparing and implementing such activities so that they can easily carry these activities into their future science classrooms. In this regard, in science teacher education program of Turkey, PSTs learn about argumentation in some courses such as science teaching and learning approaches, science teaching I and II, scientific reasoning skills and nature of science and its teaching (Council of Higher Education [CoHE], 2018). Indeed, they learn about POE technique in science teaching II and scientific reasoning skills courses. Therefore, after completing these courses, PSTs gain basic pedagogical knowledge and skills, and become familiar to argumentation. However, being familiar to argumentation is not enough for PSTs to carry out argumentative-based instructions by using discrepant events in their future science classes. To accomplish this, PSTs should also have a certain level of content knowledge and hold positive views about the possible contribution of these events for science lessons. Thus, PSTs' arguments about a discrepant event can provide information about their ability to produce qualified arguments using their content knowledge. Also, by examining their views about the use of discrepant events, their willingness to use these events in science lessons can be predicted. Based on these reasons, this study aims to determine prospective science teachers’ arguments about a discrepant event and their views on the use of these events in science education. The research questions determined for this purpose are as follows;

1) What arguments do the prospective science teachers create while predicting and explaining the results of a discrepant event?

2) What opinions do the prospective science teachers express regarding the use of discrepant events in science education?

Method

Research Design

An instrumental case study, one of the qualitative research method designs, was employed in this research. In instrumental case studies, the researcher focuses on an issue, problem or concern, and then selects a case to examine this issue (Creswell & Poth, 2018). Instrumental case study enables researchers to investigate a situation and test established perspectives on it. Moreover, similarities and differences within the boundaries that define a situation can be identified. Instrumental case study design is also used when a general understanding of a research question is needed and when it is desired to gain insight into the question by examining a particular case. In other words, researchers use a specific case to gain insight into an issue or theory (Stake, 1995). In this study, the researcher used a sample discrepant event to reveal if PSTs can craft strong arguments in the explanation of the sample event and to ascertain their thoughts about the use of these types of events in science education. In this regard, an instrumental case study design was utilized to determine whether the PSTs could explain a discrepant event using their content knowledge, the arguments PSTs develop in their attempts to explain the discrepant event, and their views on the use of discrepant events in science education.

Participants

The study group of the research is consisted of PSTs studying at third and fourth grades of science education department of a state university in Turkey. The study group was determined by using convenience sampling. In this context, PSTs in the department where the researcher works are included in the study. The first application of the activity was conducted with 33 PSTs studying in the third grade. However, the third grade prospective teachers had difficulties in explaining the discrepant event and in producing strong arguments when they studied individually. Thereupon, the research was repeated with fourth grade students. In this second application, students were asked to work in groups. The second application was carried out with 40 PSTs studying in the fourth grade in the same department. Thus, 73 PSTs participated in the study in total. The PSTs who participated in both applications of the activity have completed all physics and chemistry courses in the undergraduate programs. PSTs in the first application were coded as P1, P2, P3…, P33. In the second application, PSTs formed 10 groups and worked in groups. These groups were coded as G1, G2, G3,… G10. Although working in groups do not make a significance effect on the performance of students in tasks that require memorization and rote learning (Phelps & Damon, 1989), students working in groups perform significantly better than the ones working individually in complex and conceptual tasks (Barron, 2000) such as argumentation activities. Therefore, PSTs were asked to work in groups in the second application.

Data collection

Documents are one of the data sources in qualitative research. Examination of documents includes the use of written documents containing information about the situations desired to be studied (Yıldırım & Şimşek, 2011). In this research, written documents by PSTs were used as a data collection tool to determine the arguments of the prospective science teachers regarding a discrepant event. In this context, as a document, a paper was distributed by the researcher on which the prospective science teachers can write their predictions, observations and explanations about the discrepant event. There were four questions on the document as "What is your prediction?", "What did you observe?", "What is your explanation?" and "What do you think about use of this discrepant event in a science class? Why?" After the discrepant event was introduced, the prospective teachers wrote their predictions under the relevant title. Then, the activity was conducted and the groups of PSTs wrote their observations, explanations and their thought about the use of discrepant events under each related question with their justification. During the activity, the researcher followed and directed the individual and group discussions and enabled all prospective teachers to participate in the discussion. Then, the papers were collected and analyzed by the researcher. All the collected data in the research was in Turkish since the language of the education was Turkish and PSTs did not speak in English. Therefore, all the written material was translated into English by the researcher before the analysis.

Research Procedures

The discrepant event used in the research is adopted from a video called “discrepant balloons” (FlinnScientific, 2012). The activity in the video shared on YouTube was carried to the classroom environment by the researcher, and the same experimental setup was created in the classroom. In the discrepant balloon activity, two identical balloons were inflated to be approximately one liter and three liters respectively, and the balloons were connected to each other through a tube that allow air to pass between the two balloons. The visual of the discrepant balloon activity is given in Figure 1.

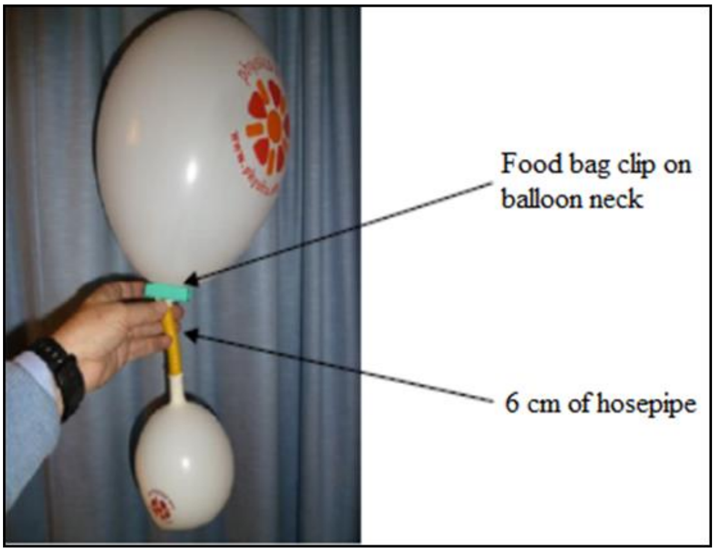


Figure 1. A picture of discrepant balloon activity (Figure is from Ješková et al., 2012)

Two balloons were connected as in Figure 1 and the airflow between the balloons was prevented. Prospective teachers were asked if the air passes from one balloon to the other when airflow is allowed, and if so, in which direction the air will flow. After PSTs wrote their predictions, airflow was allowed between the balloons. As a result of the event, it was observed that all the air in the small balloon passed into the larger one.

In the activity carried out in this research, it is expected that when airflow between the balloons is allowed, air will flow from the large balloon to the small balloon and the volumes will be equalized. However, the result of the activity was the opposite of this expectation. The reason for this situation can be explained by the amount of increase in surface areas of the balloons. When the volume of a balloon increases from one litter to two litters, its surface area increases 285 cm^2. However, when the volume of a balloon increases from 3 litters to 4 litters, the surface area of the balloon increases only 213 cm^2. The discrepancy in the behaviors of the balloons can also be experienced in daily life when a balloon is inflated. The most difficult blow into a balloon is the first one because the increase in the surface area of the balloon in the first blow is the highest, as seen in Table 1. An equal amount of air in each subsequent blow will result in a smaller increase in surface area.

Discrepant balloons activity contains a discrepancy. Therefore, it can be used to trigger argumentation process in science lessons. In this scope, the discrepant balloon activity was carried out and PSTs were asked to explain the discrepancy in this activity. The activity was carried out in two stages. In the first stage, PSTs were introduced to the discrepant event, and they were asked to express their predictions about the behaviors of balloons. Then, discrepant balloons activity was demonstrated. After that, PSTs were asked to write their observations and explanations about the results of the activity. The predictions and explanations of PST were analyzed by the researcher. As a result of the analysis, it was seen that there were important deficiencies and inaccuracies in the explanations of the PSTs, and it was decided to re-organize the activity with a different group of PSTs by giving some supportive information.

Before the second demonstration, the researcher set a session with PSTs to introduce the discrepant balloon activity, Toulmin Argumentation Pattern (2003), and POE technique. In this session, the researcher introduced Toulmin model, explained its components and gave examples of strong and weak arguments from science education topics on the smart board in a science laboratory. Also, each step of POE technique was introduced and reminded to PSTs in this session. PSTs were asked to craft convincing, persuasive and strong arguments composing of justification, qualifier and multiple rebuttals as in the examples that researcher showed. In other words, they were asked to justify their claims and create strong arguments by using their content knowledge. Before the discrepant balloon activity was conducted by the researcher, PSTs were asked to create their arguments regarding their predictions about the activity.

While the PSTs were making predictions about the discrepant event in the second application, firstly, they held small group discussions within their groups. In these small group discussions, they were asked to create an argument about what is going to happen when the airflow is allowed between balloons and why. They were asked to defend their arguments to group members. After consensus was reached in small groups, all groups explained their arguments regarding their predictions to the researcher and other groups. After listening to the arguments of all groups, the researcher performed the demonstration. Besides the demonstration, some information about the volumes, radii, and surface areas of the balloons (Table 1) (FlinnScientific, 2012) and a graph showing the radius-pressure change of a balloon (Figure 2) was shared with PSTs to support them with some extra information. The information provided to PSTs about the activity was adapted from the research by Ješková et al. (2012). PSTs were asked to consider both their observations and the given information in their explanations about the discrepant event. Then, PSTs were allowed to re-evaluate the arguments they formed in the prediction phase. Thereby, they started a new argumentation process within their groups based on their observation and the information given to them. At this step, they either developed their previous argument or formed a new one.

Small group discussions to explain the observed discrepancy continued until a consensus was reached in each group. Then, the researcher listened to the groups' arguments one by one and started whole class discussions between the groups. After enough time was given to the discussions between all groups, the groups were asked to express in written form whether their ideas have changed or not. Lastly, the PSTs were asked about the possible contributions of using the discrepant events in science education, and they were asked to write their opinions. The activity ended after PSTs wrote their opinions.

Table 1. Information about balloons

Volume (cm^3)	Radii (cm)	Surface Area (cm^2)
0	0	0
1000	6.20	483
2000	7.82	768
3000	8.95	1006
4000	9.85	1219

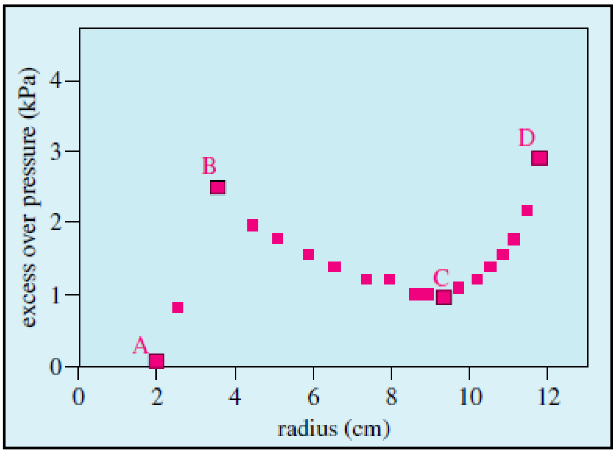


Figure 2. Experimental result of pressure versus radius for a balloon (Ješková et al., 2012)

Data Analysis

Descriptive analysis was used in the analysis of the data of this research. Descriptive analysis includes summarizing and interpreting qualitative data collected by different methods based on predetermined themes. In descriptive analysis, direct quotations can be frequently used in order to present the feelings and thoughts of the people with whom the researcher directly interacted. The purpose of the descriptive analysis method is to transfer the obtained data to the reader in a collective and interpreted manner. Descriptive analysis takes place in four stages. These stages are creating the framework required for descriptive analysis, processing the data, identifying the findings and interpreting them (Altunışık et al. 2010; Yıldırım & Şimşek, 2011). In the first stage, the limits of data analysis were drawn by the researcher based on the research problems and the collected data. In this study, data about the PSTs' arguments regarding the discrepant event and opinions about the use of discrepant events in science education were obtained. In the second stage, the researcher put the data into a framework that was previously created. The data for the first research question was listed under the themes determined in accordance with the phases of POE technique. Sub-themes were created based on the components of an argument and were separately given under the relevant themes. The data regarding the second research question were coded as labelling the participants' views. In the third stage, the findings put into order by the researcher are explained. Direct quotations can be given where needed in this stage. In this study, findings were presented with direct quotations of PSTs. In the fourth stage, researchers explain the findings and interpret the meaning by establishing a link between them. In this study, logical inferences in the scope of the fourth stage were made in the conclusion and discussion section.

Validity and reliability

In the research expert approval was obtained before using the POE form to ensure the credibility of the research. In this regard , opinions of one expert at educational sciences were taken regarding the prepared interview form. Also, two experts at science education examined the discrepant event activity and approved the discrepancy and explanation suggested by the researcher. Also, research design is approved by these experts. The research model, study group, data collection tools and procedures, data analysis and were organized are described in detail in method section to ensure transferability. In addition, the participants of the study were prospective science teachers who can contribute to the purpose of the study. However, relatively small number of participants can be shown as a limitation for transferability. In terms of dependability (internal reliability), findings of the research were presented without comment, and data loss was prevented since the data is collected through written documents. In addition, the codes created by the researcher is examined and approved by an external researcher who is experienced in qualitative research for confirmability (external reliability). Finally, the results are appropriately discussed in accordance with the present literature. Therefore, the confirmability of the study was aimed to be increased.

Findings and Discussions

Findings Regarding PSTs' Arguments in First Application

The discrepant balloons activity used in the research was first demonstrated to 33 PSTs. In this context, firstly, findings related to the predictions, observations and explanations of prospective teachers about discrepant event are given in Table 2.

Table 2. Codes and themes from the first application

Theme	Codes	Participants
Prediction	Both balloons deflate	P2, P31
	Smaller balloon deflate completely, bigger balloon deflate partially	P1, P3
	The bigger balloon deflates some, the smaller balloon inflates some and they come to equal volume.	P4, P5, P6, P9, P15, P16, P17, P18, P19, P20, P21, P22, P23, P24, P25, P26, P27, P28, P29, P30, P32, P33
	The smaller balloon deflates completely, the bigger balloon inflates.	P7, P8, P11, P12, P13
	Volumes remain unchanged	P10, P14
Observation	The smaller balloon deflated completely, the bigger balloon inflated.	All of the participants
Explanation	Density difference	P1, P2, P3, P7, P10, P28, P29
	Low gas pressure in the small balloon	P1, P4, P25, P31

	Low gas pressure in the bigger balloon	P8, P11, P13, P15, P16, P17, P18, P19, P20, P21, P22, P24, P26, P27, P30
	Small amount of gas in the small balloon	P5, P6, P18
	Volume difference	P9, P10, P12,
	Expansion	P14

* P23, P32 and P33 could not make any explanation regarding the cause of the event. In addition, some participants' (P1, P10 and P18) explanations could be collected under two different codes.

In the study, the predictions of PSTs regarding discrepant balloons activity were determined. Two participants (P2 and P31) stated that both balloons would be completely deflated. Among them, P31 expressed her thoughts as "*If the clip between two balloons is opened, both balloons will be deflated*". Among PSTs who think that the small balloon will be completely deflated and the large balloon will partially deflate, P1 stated that "*while the small balloon goes out, larger balloon also goes out partially. So a big part of larger balloon goes out.*" On the other hand, most of the PSTs predicted that the big balloon would deflate a little and the small balloon would swell a little and eventually they come to equal volumes. P6's prediction is as fallows "*Air flow from bigger balloon to small one. It continues until the volumes of two balloons are equalized*." Some of the PSTs made the correct prediction about the activity. P8 who predicted that the small balloon would completely deflate and the large balloon would swell stated her prediction as;

> "*The air passes from the small balloon to the big balloon. Since the volume of the small balloon is low, its pressure is high. Due to the high pressure, the air passes from the small balloon to the big one. So the big balloon will inflate more.*"

The PSTs coded as P10 and P14, stated that when air passage between the balloons is allowed, there will be no change in the size of the balloons. P14 stated his opinion as, "*I think balloons remain the same. There will be no change. In order for diffusion to occur, the volume of the small balloon must increase.*"

Discrepant balloons activity was held after PSTs expressed their predictions in writing. All PSTs stated in their observations that the small balloon completely deflated and the large balloon swelled more. After the prospective science teachers explained their observations, they were asked to explain the reason of the situation.

When the answers of the PSTs are examined, it is seen that they emphasized certain concepts such as density, gas pressure, gas amount and expansion in order to explain the behaviors of the balloons. According to P1 who thinks that the air in the small balloon passed to larger one due to the difference in density. To reveal her ideas, she stated that "*the small balloon went out, the big balloon swelled because the gas density in the small balloon was less than the gas density in the large one*". P31 emphasized to the low gas pressure in the small balloon and stated that;

> "*Since the gas pressure in the small balloon is less, the large balloon with more air does not go out. The gas in the small balloon passes into the big one. It goes out as the gas pressure is low in the small balloon*."

Most of the PSTs explained the discrepant event with the large gas pressure in the big balloon. For example, P13 expressed his opinion on this issue as follows;

> "*The small balloon completely deflated because its volume was small. (Therefore) the pressure was high… The volume of the big balloon was large. So, the pressure was low. Air flows from high pressure to the low."*

Three PSTs explained the discrepancy in the activity with the amount of gas in the balloons. P6 explained his opinion as "*I think that the small balloon deflated because there is less amount of gas in it*". Explaining the discrepant event with the different volumes of the balloons, P10 thoughts are as fallows; "*since the volume of the small balloon is smaller than the large one, there was a gas transition from the small balloon to the large. I mean because of the difference in volume*." Explaining the discrepant event through the concept of expansion, P14 explained her thought as "since the large balloon has expanded more, the pressure of the small balloon has effected easily on the large balloon".

In the first application, while the PSTs' arguments about their predictions consisted of only the claim. In the second application, with the guidance of the researcher, the PSTs were able to create more developed arguments consisting of the claim and warrants. Although PSTs were able to create arguments consisting of the claim and warrants, argument components such as rebuttal and backing were not found in any argument. From this aspect, it can be said that the arguments created in the first application are quite weak. This means that PSTs' content knowledge regarding the pressure, diffusion, and density concepts is superficial or they cannot use their content knowledge to explain a phenomenon. In the literature, there are results reporting that PSTs have difficulties in producing high-level arguments containing rebuttal and backing (Erduran et al., 2004; Hiğde & Aktamış, 2017). Regarding their content knowledge, for example, it is seen that PSTs have difficulty in explaining the relationship between pressure, volume and density. By associating pressure only with volume and ignoring other variables such as temperature and the amount of air, they reach the proposition that pressure increases as the volume decreases. Similarly, PSTs in G8 associated pressure only with the amount of air in the balloons and neglected the volume. In one group, students' justification was not even scientific. The PSTs in this group stated that there is a balance in nature and therefore, balloons must behave in line with the balance of the nature, which shows that PSTs in G9 justified their claim with a non-scientific "argument". The justification of this group was teleological rather than scientific. This shows that, PSTs have not developed conceptual understanding in pressure concept. The most important reason for this situation is that PSTs learn about physics and chemistry concepts with an approach far from daily life practices. More precisely, they tend to learn these concepts by memorizing. Therefore, the insufficiency in content knowledge is an obstacle for PSTs to produce strong scientific arguments. Supporting these results, it is stated in the literature that the pedagogical competencies of teachers highly depend on their content knowledge (Canbazoğlu et al., 2010; Magnusson et al., 1999). Besides, there are also supporting results in the literature that teachers tend to ask more superficial questions while teaching the subjects on which their content knowledge is insufficient (Carlsen, 1999).

Findings Regarding PSTs' Arguments in Second Application

The discrepant balloons activity was demonstrated to a group of PSTs. In this context, findings related to the predictions, observations, and explanations of prospective teachers about the discrepant event are given in Table 3.

Table 3. Codes and themes from the data

Theme	Sub-theme	Code	Groups
Prediction	Claim	The bigger balloon deflates some, the smaller balloon inflates some and they come to equal volume.	G1, G5, G6, G7, G8, G9, G10
		The smaller balloon deflates completely, the bigger balloon inflates.	G2, G3, G4
	Warrant	Diffusion	G1, G5
		The internal pressure of the small balloon is higher.	G4,
		The internal pressure of the large balloon is higher.	G6, G7
		There is more air in the big balloon.	G8
		There is a balance in nature.	G9
		Gases pass from high density to low.	G10

Observation		The smaller balloon deflated completely, the bigger balloon inflated.	All of the groups
Explanation	Warrant	The pressure decreases as the surface area and radius of the balloon increase	G1
		The pressure increases as the surface area decreases	G2, G4
		External pressure decreases as the membrane gets thinner	G2
		The balloons will swell harder when its membrane is thick	G3
		As the balloon swells, its elasticity changes	G6
		The membrane of the small balloon exerts less force on the air in it	G7
		The membrane of the small balloon is thicker	G8
		The big balloon flexes more easily than the small one	G9
		The pressure decreases as the radius increases	G1, G10

In the research, PSTs groups were asked to write their predictions with their justifications. When the predictions of PSTs groups were examined, two codes regarding their claims were formed. Some groups of the PSTs stated that there would be an air transition from the big balloon to the small one until the volumes of the balloons were equalized. For example, PSTs in group six (G6) expressed their views on this issue as "*Air flows from the larger balloon to the small balloon, and air passes until their volumes are equalized.*"

Similarly, PSTs in G1 stated that "*there will be air flow from large to the small balloon. When they come to the same size, air flow stops.*" On the other hand, some groups of the PSTs expressed that when the clip between the balloons was opened, the small balloon would completely deflate and the large balloon would inflate a little more. For example, the PSTs in the fourth group (G4) revealed their thoughts as "*air passes from small balloon to large balloon.*" Similarly, PSTs in G3 stated that "air pass from small balloon to the large one. Small balloon deflates completely."

In the research, PSTs in groups were asked to justify their predictions. It is noteworthy that PSTs emphasized diffusion, pressure, density, and balance concepts in their justifications. For example, PSTs in the G5 who thought that air would pass from the big balloon to the small balloon due to the diffusion justified their thoughts as "*they will be equalized in volume because the air passes from high density to low due to the diffusion rule*." PSTs in G4, G6, and G7 used the pressure concept to justify their claims. The PSTs in the G4 thought that the internal gas pressure in the small balloon was higher. They expressed their thoughts as "*The volume of the small balloon is low. Therefore, the internal pressure is high. Large balloon has a high volume and internal pressure is low*."

The PSTs in G6 thought that the internal pressure of the big balloon was higher, and they explained their thoughts as "*because the air pressure is too high in the larger balloon, air passes into the small balloon. The air passage continues until their pressures are equalized*." The PSTs in G8 justified their claim by referring to the amount of the air in the balloons. They uttered that "*air passes from larger balloon to the small one because the amount of air in the larger balloon is much more than that of the small balloon*." The PSTs in the G9 put forward a metaphysical justification for their claims. Their justification was *"... the air transition continues until the balloons reach equilibrium because the air is a fluid and there is a balance in nature*."

After their argument in the prediction phase, all groups observed that all the air in the small balloon passed to the larger balloon. Therefore, all groups reported this observation commonly. In this context, there was no difference in terms of the observations of the groups.

In the third stage of POE, PSTs in groups were asked to explain the reasons for the results of the discrepant event. In other words, they were asked to write why all the air in the small balloon passed into the large balloon. When the answers of PSTs groups were examined, it was seen in their explanation that, they emphasized the surface area, radii, and membrane factors. For example, PSTs in G1, who emphasized the surface area and radius, explained their thoughts as "*the large balloon has more surface area and radius. This means the pressure of the big balloon is smaller and the pressure of the small balloon is higher. Air passes from high pressure to low pressure*."

PSTs in G2 highlighted the surface area and noted that "*since the small balloon has a small surface area, it has a higher pressure. Therefore, air passed from small balloon to the large one*". PSTs in G2 also made a contradictory justification. They stated that "*external air pressure is equal for both balloons…as the membrane of the balloon gets thinner, external air pressure decreases*." On the other hand, it was seen that PSTs in many groups (G2, G3, G6, G7, G8 and G9) emphasized the membrane of the balloon to explain the discrepant event. For example, the PSTs in G3 explained their opinions as "*when we try to blow up a balloon, it is difficult to inflate it because its membrane is thick*." Also, PSTs in G8 asserted that "*since the membrane of the small balloon is thicker, it does not tend to swell*." Similarly, PSTs in G6 explained the situation as follows;

"*….the reason for this may be that the flexibility of the balloon changes over time when we blow into the balloon. … Because the small balloon is less flexible and the large balloon is more flexible, the air in the small balloon can easily pass into the large balloon*."

In the second application, it was seen that the PSTs' arguments regarding the results of the activity were relatively more qualified and the power of explanation was higher. Also, arguments regarding the results of the activity were relatively more qualified and the power of explanation was higher compared to the arguments in first application. The main reasons for the increase of the argument quality may be additional information provided and working as a group. Thanks to peer instruction and peer inquiry in group work, PSTs were able to see the deficiencies or errors in their own ideas, and they were able to look critically. In this way, they were able to generate stronger arguments. In this regard, groups referred to the information such as surface area and radii given them. Supporting, previous studies reported that providing supportive information to students enables students to produce more qualified arguments (Akbayrak & Namdar, 2019; Schworm & Renkl, 2007). In addition, PSTs explained the results of the activity by making logical inferences from the available data. Therefore, it can be said that especially small group discussions helped PSTs to think analytically. In the literature, there are studies supporting the results in the research that peer instruction and peer inquiry improve conceptual learning (Kızkapan & Bektaş, 2021; Mazur, 1997). Based on these results, it can be said that activities that will enable individuals to produce qualified arguments with very cheap materials and correct pedagogical methods can be held in science classes.

Findings Regarding the Use of the Discrepant Event in Science Classrooms

Within the scope of the second sub-problem of the research, the PSTs' opinions about the use of the discrepant events in science classroom were asked in the second application. PSTs expressed their thoughts as a group after small group discussions about the possible effects, positive and negative aspects of using the discrepant balloon activity in science classroom. The codes obtained from the analysis of PSTs' answers are given in Table 4.

Table 4. PSTs' opinions regarding the use of the discrepant event

Codes	Groups
21st-century skills	G1, G2, G4, G5, G8, G10
Eliminating misconceptions	G3, G4, G9
Cooperative learning	G2, G4, G8

Research -Inquiry- Curiosity	G1, G4, G5, G7, G8
Conceptual and permanent learning	G2, G3, G6, G7, G8, G9
Attitude and motivation	G1, G4, G8
Draw attention	G4, G7
Psychomotor skills	G7
Science process skills	G1, G2, G7

The opinions of some groups regarding the 21st-century skills code about the benefits of using the discrepant event in science education are as "*discrepant events should be used in science classes because they improve students' creative thinking skills, supports group work and collaborative work*" (G8), "*such activities enable students to come up with problem-solving, analytical thinking, critical thinking, creative ideas*" (G4), "*such experiments should be carried out because these experiments make the student active participants of the lessons and improve their innovation skills*" (G2).

The groups thinking that discrepant events would eliminate the misconception of the students expressed their opinions as "*these activities enable students to consolidate the concepts and eliminate misconceptions*" (G9), "*such activities can be used to arouse students' curiosity and eliminate misconceptions*" (G4).

Three groups have emphasized cooperative learning. PSTs in one of these groups expressed their thoughts as "*such activities should be used because they enhance collaboration in group work*" (G8).

Five groups stated that discrepant events would improve students' curiosity and some skills such as research and inquiry. Thoughts of some of these groups' are as "*such activities develop a sense of curiosity in students and improve research skills*" (G1), "*these types of events can be used in lessons. These activities improve creative thinking. It awakens a sense of curiosity*" (G5).

The majority of groups stated that discrepant events would provide conceptual and permanent learning. Some groups' views on this regard were as "*these activities need to be used in science lessons because students can learn conceptually because it is shown as an application*" (G6), "*such activities should definitely be included in science teaching. They enable students to learn meaningfully and permanently by doing and living*" (G7).

Some groups stated that discrepant events would increase students' attitude and motivation towards science lessons. Thoughts of some of these groups were as "*with such* *activities, students' willingness to learn increases*" (G4), "*we should do these activities because they increase students' positive attitudes towards science lessons*" (G8).

PSTs who thought that discrepant events could be used to draw attention expressed their thoughts as "*such activities should be included in science lessons because they draw students' attention*" (G7), "*such experiments can be used to draw students' attention to the lesson, to arouse curiosity*" (G4).

The PSTs in G7 stated that discrepant events could be used to develop psychomotor skills. Group members expressed their opinions as "*such activities should be included in science education. In this way, students' psychomotor skills can be developed*" (G7).

Lastly, PST groups stated that discrepant events would improve students' scientific process skills. Some groups' thoughts on this were as "*students' prediction, observation, and interpretation skills can be improved*" (G1), "*we must provide students with scientific process skills. Therefore, such activities should be used*" (G2).

In the current research, regarding the possible benefits of using discrepant events in science education, all of the groups stated that it is beneficial to use discrepant events in science education. At this point, PSTs mostly emphasized that discrepant events can improve students' 21st-century skills, provide meaningful and permanent learning, and improve students' curiosity, research, and inquiry skills. Discrepant events are quite powerful tools that can be used to arouse interest, learn about misconceptions, and develop higher-order thinking skills (González - Espada et al., 2010). Therefore, PSTs' expression that discrepant events will contribute to science education is an indication that they will include such activities in their classrooms in the future. In this way, they will be able to create learning environments in the future where their students can produce more qualified arguments, develop higher-level thinking skills, and evaluate the events from a critical and analytical perspective.

Conclusions and Suggestions

Results of this research showed that PSTs could not make high-level arguments to explain a discrepant event by using their physics and chemistry knowledge in the prediction phase. However, when they are provided with additional information that they can use to explain the discrepancy and well facilitated during argumentation, the quality and explanatory power of the arguments increased slightly. Finally, it was seen that PSTs have the opinions that using discrepant events would be beneficial in science education.

Based on the results, it can be suggested that PSTs should be exposed to different discrepant events or situations that create different cognitive contradictions during their university education. In other words, teaching environments where PSTs can create more qualified arguments for discrepant events should be created. In teaching method courses within the scope of teacher education programs in Turkey, the discrepant event should be used as a productive context in teaching some certain methods such as argumentation or inquiry based learning. Also, results of the study revealed that PSTs have difficulty in using their content knowledge to explain a scientific phenomenon. In order to support learners from all levels including university education, constructivist and learner centered teaching approaches should be adopted in science lessons. In this regard, students should face to real life problems they need to solve using their content knowledge. In this regard, use of discrepant events can be helpful. In addition, as a limitation of the current research, only one discrepant event is used to asses PSTs arguments. Additional researches implementing different discrepant events from different disciples can be conducted. Finally, it is recommended to conduct researches with students from different educational levels to develop competencies such as analytical and critical thinking, conceptual learning, and attitudes by using discrepant events.

Fen Bilgisi Öğretmen Adaylarının Sezgiye Ters Bir Olaya İlişkin Argümanları ve Bu Tür Olayların Fen Eğitiminde Kullanılmasına İlişkin Düşünceleri

Özet:

Bu çalışmanın amacı, fen bilgisi öğretmen adaylarının (FBÖA) sezgiye ters bir olaya ilişkin argümanlarını ve bu tür çelişkili olayların fen eğitiminde kullanımına ilişkin görüşlerini belirlemektir. Araştırmada nitel araştırma yöntemi desenlerinden araçsal durum çalışması benimsenmiştir. Araştırmanın çalışma grubunu Türkiye'de İç Anadolu Bölgesinde bulunan bir devlet üniversitesinin fen bilgisi öğretmenliği bölümünde okuyan 73 öğretmen adayı oluşturmaktadır. Araştırmanın verileri yazılı dokümanlar kullanılarak toplanmıştır. Araştırmada toplanan veriler betimsel analiz kullanılarak analiz edilmiştir. Analizler, fen bilgisi öğretmen adaylarının sezgiye ters olayla ilgili bireysel argümanlarının zayıf ve çelişkiyi açıklamakta yetersiz olduğunu göstermiştir. Fen bilgisi öğretmen adaylarına ek bilgiler sağlandığında, etkili bir şekilde işbirliği yapmaya teşvik edildiğinde ve grup olarak çalışmalarına izin verildiğinde ise argümanların kalitesi ve açıklama gücü artmıştır. Ancak öğretmen adaylarının ne bireysel ne de grup olarak çalıştıklarında üst düzey argümanlar üretemedikleri görülmüştür.

Anahtar kelimeler: sezgiye ters olaylar, fen eğitimi, argümantasyon, kimya eğitimi

References

Akbayrak, K., & Namdar, B. (2019) An argumentation activity for third-grade students: objects in the plates, *Science Activities, 56*(1), 1-10. https://doi.org/10.1080/00368121.2019.1600464

Altunışık, R., Coşkun, R., Bayraktaroğlu, S., & Yıldırım, E. (2010). *Sosyal bilimlerde araştırma yöntemleri: SPSS uygulamalı [Research methods in social sciences: SPSS applications].* Sakarya Publishing.

Anderson, C.. & Smith, E. (1983). Teacher behavior associated with conceptual learning. In *Proceedings of the Annual Meeting of the American 'Educational Research Association*, Montreal. Canada.

Appleton, K. (1996). Students' responses during discrepant event science lessons. In *Proceedings of the annual meeting of the National Association for Research in Science Teaching*, (ERIC Document Reproduction Service No. ED 393 696)

Bakioğlu, B., & Çevik, M. (2020). Views of science teachers on distance education during the COVID-19 pandemic. *Turkish Studies, 15*(4), 109-129. https://dx.doi.org/10.7827/TurkishStudies.43502

Barron, B. (2000). Problem solving in video-based micro worlds: Collaboration and individual outcomes of high achieving sixth grade students. *Journal of Educational Psychology, 92*(2), 391 – 398. https://psycnet.apa.org/doi/10.1037/0022-0663.92.2.391

Blikstein, P., Fuhrmann, T., & Salehi, S. (2016). Using the bifocal modeling framework to resolve "Discrepant Events" between physical experiments and virtual models in biology. *Journal of Science Education and Technology*, *25*(4), 513-526. https://doi.org/10.1007/s10956-016-9623-7

Canbazoğlu, S., Demirelli, H., & Kavak, N. (2010). Investigation of the relationship between pre-service science teachers' subject matter knowledge and pedagogical content knowledge regarding the particulate nature of matter. *Elementary Education Online*, *9*(1), 275-291. https://dergipark.org.tr/tr/download/article-file/90808

Carlsen, W. S. (1999). Domains of teacher education. In J. Gess-Newsome and N.G. Lederman (Eds.), *Examining Pedagogical Content Knowledge*, (133-144). Kluwer Academic Publishers.

CoHE, (2018). New teacher training undergraduate programs. Retrieved from https://www.yok.gov.tr/kurumsal/idari-birimler/egitim-ogretim-dairesi/yeni-ogretmen-yetistirme-lisans-programlari

Creswell, J. W., & Poth, C. N. (2018). *Qualitative inquiry and research design: Choosing among five approaches*. Sage publications.

de Oliveira, M. H. A., & Fischer, R. (2017). Ciênsação: gaining a feeling for sciences. *Physics Education*, *52*(2), 025011.

Erduran, S., Simon, S., & Osborne, J. (2004). TAPping into argumentation: Developments in the application of Toulmin's argument pattern for studying science discourse. *Science education*, *88*(6), 915-933. https://doi.org/10.1002/sce.20012

Fitriani, A., Zubaidah, S., Susilo, H., & Al Muhdhar, M. H. I. (2020). The effects of integrated problem-based learning, predict, observe, explain on problem-solving skills and self-efficacy. *Eurasian Journal of Educational Research*, *20*(85), 45-64. https://doi.org/10.14689/ejer.2020.85.3

FlinnScientific. (2012). *Discrepant balloons*. Retrieved from https://www.youtube.com/watch?v=W12k1sTw-do.

González-Espada, W. J., Birriel, J., & Birriel, I. (2010). Discrepant events: A challenge to students' intuition. *The Physics Teacher*, *48*(8), 508-511. https://doi.org/10.1119/1.3502499

Harteis, C., Goller, M., and Caruso, C. (2020). Conceptual change in the face of digitalization: challenges for workplaces and workplace learning. *Frontiers in Education. 5*(1), 1-10. https://doi: 10.3389/feduc.2020.00001

Hewson, P. W., & Hewson, M. G. B. (1984). The role of conceptual conflict in conceptual change and the design of science instruction. *Instructional Science*, *13*(1), 1-13. https://doi.org/10.1007/BF00051837

Hiğde, E., & Aktamış, H. (2017). Examination of pre-service science teachers' argumentation-based science lessons: case study. *Elementary Education Online*, *16*(1), 89-113. http://dx.doi.org/http://dx.doi.org/10.17051/io.2017.79802

Hong, J. C., Hsiao, H. S., Chen, P. H., Lu, C. C., Tai, K. H., & Tsai, C. R. (2021). Critical attitude and ability associated with students' self-confidence and attitude toward "predict-observe-explain" online science inquiry learning. *Computers & Education*, *166*, 104172.

Ješková, Z., Featonby, D., & Feková, V. (2012). Balloons revisited. *Physics Education*, *47*(4), 392-398. https://doi.org/10.1088/0031-9120/47/4/392

Karamustafaoğlu, S., & Mamlok-Naaman, R. (2015). Understanding electrochemistry concepts using the predict-observe-explain strategy. *Eurasia Journal of Mathematics,* *Science and Technology Education*, *11*(5), 923-936. https://doi.org/10.12973/eurasia.2015.1364a

Kavogli, Z. (1992). Discrepant events: An alternative teaching process. *Science Education International*, *3*(3), 10-13.

Kızkapan, O. & Bektaş, O. (2021) Enhancing seventh-grade students' academic achievement through epistemologically enriched argumentation instruction, *International Journal of Science Education*, *43* (10), 1600-1617, https://doi.org/10.1080/09500693.2021.1923082

Limón, M. (2001). On the cognitive conflict as an instructional strategy for conceptual change: A critical appraisal. *Learning and instruction*, *11*(4-5), 357-380. https://doi.org/10.1016/S0959-4752(00)00037-2

Magnusson, S., Krajcik, J. & Borko, H. (1999). Nature, sources, and development of pedagogical content knowledge for science teaching. In J. Gess-Newsome and N.G. Lederman (Eds.), *Examining Pedagogical Content Knowledge* (pp. 95–132). Dordrecht, Kluwer Academic Publishers.

Mancuso, V. J. (2010). *Using discrepant events in science demonstrations to promote student engagement in scientific investigations: An action research study.* Doctorate dissertation, University of Rochester, Rochester, New York

Mason, D., Griffith, W. F., Hogue, S. E., Holley, K., & Hunter, K. (2004). Discrepant event: The great bowling ball float-off. *Journal of Chemical Education*, 81, 1309-1312. https://doi.org/10.1021/ed081p1309

Mazur, E. (1997). *Peer Instruction: a User's Manual*. Prentice-Hall.

Nadelson, L. S., Heddy, B. C., Jones, S., Taasoobshirazi, G., & Johnson, M. L. (2018). Conceptual change in science teaching and learning: Introducing the dynamic model of conceptual change. *International Journal of Educational Psychology*, **7**(2), 151-195. https://doi.org/10.2196/resprot.2372

O'Brien, T., Stannard, C., & Telesca, A. (1994). A baker's dozen of discrepantly dense demos, *Science Scope*, *18*, 35–38.

Phelps, E., & Damon, W. (1989). Problem solving with equals: Peer collaboration as a context for learning mathematics and spatial concepts. *Journal of Educational Psychology*, *81*(4), 639-646. https://psycnet.apa.org/doi/10.1037/0022-0663.81.4.639

Posner, G. J., Strike, K. A., Hewson, P. W., & Gertzog, W. A. (1982). Accommodation of a scientific conception: Toward a theory of conceptual change. *Science Education*, *66*(2), 211-227.

Sampson, V., & Clark, D. (2009). The impact of collaboration on the outcomes of scientific argumentation. *Science Education*, *93*(3), 448–484. https://doi.org/10.1002/sce.20306.

Sampson, V., Grooms, J., & Walker, J. P. (2011). Argument-Driven Inquiry as a way to help students learn how to participate in scientific argumentation and craft written arguments: An exploratory study. *Science Education*, *95*(2), 217-257. https://doi.org/10.1002/sce.20421

Sarah, S., Khanif, A., & Saputra, A. T. (2021). The effectiveness of POE (Predict-Observe-Explain) learning model for ımproving student analytical skills. *JIPF (Jurnal Ilmu Pendidikan Fisika)*, *6*(1), 23-29.

Shemwell, J. T., & Furtak, E. M. (2010). Science classroom discussion as scientific argumentation: A study of conceptually rich (and poor) student talk. *Educational Assessment*, *15*(3-4), 222-250. https://doi.org/10.1080/10627197.2010.530563

Stake, R. E. (1995). *The art of case study research*. Sage Publications.

Schworm, S., & Renkl. A. (2007). Learning argumentation skills through the use of prompts for self-explaining examples. *Journal of Educational Psychology 99*(2), 285–96. https://psycnet.apa.org/doi/10.1037/0022-0663.99.2.285

Senemoğlu, N. (2010). *Development, learning, and instruction: From theory to application (17th edition).* Pegem Publishing.

Sinatra, G. M., Kienhues, D., & Hofer, B. K. (2014). Addressing challenges to public understanding of science: Epistemic cognition, motivated reasoning, and conceptual change. *Educational Psychologist*, *49*(2), 123– 138. https://doi.org/10.1080/00461520.2014.916216

Suprapto, N. (2020). Do we experience misconceptions?: An ontological review of misconceptions in science. *Studies in Philosophy of Science and Education*, *1*(2), 50-55. https://doi.org/10.46627/sipose

Taşçı, S. (2021). Evaluation of emergency distance language education: Perspectıves of elt students. *Nevşehir Hacı Bektaş Veli University Journal of Social Sciences Institute*, *11*(1), 286–300. https://doi.org/10.30783/nevsosbilen.877657

Toulmin, S. E. (2003). *The uses of argument*. Cambridge university press.

Vaughn, A. R., Brown, R. D., & Johnson, M. L. (2020). Understanding conceptual change and science learning through educational neuroscience. *Mind, Brain, and Education, 14*(2), 82-93.

Wright, E. L., & Govindarajan, G. (1992). Stirring the biology teaching pot with discrepant events. *The American Biology Teacher*, *54*(4), 205-210. https://www.jstor.org/stable/4449456

Yıldırım, A & Şimşek, H., (2011). *Qualitative research methods in social sciences*. Seçkin Publishing

Necatibey Eğitim Fakültesi Elektronik Fen ve Matematik Eğitimi Dergisi
Cilt 16, Sayı 2, Aralık 2022, sayfa 406-427. ISSN: 1307-6086
Necatibey Faculty of Education Electronic Journal of Science and Mathematics Education
Vol. 16, Issue 2, December 2022, pp. 406-427. ISSN: 1307-6086

Research Article

Analysis of the Experiments in the Chemistry Textbooks in terms of the Laboratory Study Goal of the 2018 Chemistry Curriculum and Comparison with the Case of Science High School

Canan NAKİBOĞLU

Balıkesir University, Necatibey Education Faculty, Department of Chemistry Education, Balıkesir/Türkiye, canan@balikesir.edu.tr, https://orcid.org/0000-0002-7292-9690

Received : 24.10.2022 Accepted : 8.12.2022
Doi: 10.17522/balikesirnef.1193998

Abstract – This study aims to examine to what extent the goal of the 2018 Chemistry Curriculum for experimental studies overlaps with the acquisitions of the curriculum and to what extent upper-secondary school chemistry textbooks (high school) are prepared to meet this purpose. The study also seeks to assess the extent to which the activities in the chemistry textbooks are written to meet the four dimensions (gaining data by experimenting, inferring using data, interpretation, and generalization) of the 2018 Chemistry Curriculum. Finally, it is intended to compare this situation with the 2018 Science High School Chemistry Curriculum and determine the similarities and differences. For these purposes, the acquisitions including the experimental studies in the 2018 Chemistry Curriculum were determined at first. Then, the experiments in the 9th, 10th, 11^{th,} and 12th grade chemistry textbooks were analyzed. Two ways were used for textbook analysis. In the first stage, it was defined to what extent the acquisitions in the 2018 Chemistry Curriculum and the experiments in the textbooks overlapped and how the experiments were distributed according to the units. In the second stage, the experiments in the textbooks were analyzed according to the four dimensions of the 2018 Chemistry Curriculum using a rubric. It was concluded that the acquisitions that include explanations about experimental studies are approximately 16% of the total acquisitions of the 2018 Chemistry Curriculum. Additionally, as a result of this analysis for SPSs, it is found that the experiments mostly focus on the basic SSBs, and they focus especially on "observation" and "measurement" skills.

Keywords: 2018 Chemistry Curriculum, upper-secondary school chemistry textbook, experimental study.

Corresponding author: canan@balikesir.edu.tr

Introduction

It is essential to provide teaching experience that will enable students to better understand science subjects so that countries do not fall behind in scientific and technological developments, and to raise individuals who can produce knowledge and technology (Ayas, 1995; Zorluoğlu et al., 2016). For this reason, the approaches used in the development of secondary education chemistry curricula are very important. In recent years, it is seen that scientific process skills, research and inquiry-based activities, and acquisitions that will improve students' scientific literacy are included in the curricula developed both in the world and in Turkey (Ministry of National Education (MoNE), 2018a). Mattheis and Nakayama (1988) emphasized that the most important goals of science education are the development of laboratory skills and scientific process skills necessary for the process and applications of science, and the understanding of scientific information that is the product of these, and they argued that laboratory work has an important role in achieving these goals. Besides conducting laboratory studies within the scope of upper-secondary school chemistry courses is crucial and provides students with the opportunity to conduct experiments and understand the nature of research (Nakiboğlu & Şen, 2020). In addition, laboratory studies contribute to students in different dimensions such as problem-solving and scientific process skills, understanding the nature of science, and increasing scientific knowledge. While doing experiments, students can establish the relationship between their experiment and the scientific knowledge behind the experiment (Uzezi & Zainab, 2017).

Besides, the laboratory studies during chemistry lessons and interpretation of findings and inferences during trials can contribute to the students' meaningful learning of chemistry concepts and subjects. So, Ayas (1995) stated that the basic philosophy of using laboratories in teaching is to try and observe the relationships between phenomena and/or concepts and to make inferences based on them. Developing only psychomotor skills during experimentation in a laboratory environment is not sufficient to achieve the goals expected from laboratory studies, and it is extremely important that students are designed to develop high-level thinking skills in order to achieve the desired efficiency from experimental studies. Thus, laboratory studies are included among the general objectives and acquisitions of upper-secondary chemistry education programs.

Demir and Nakiboğlu (2021) stated that curricula predict the knowledge, skills, and behaviors that individuals should have in the short and long term, and thus form a roadmap for gaining these knowledge, skills, and behaviors. Depending on this roadmap, chemistry teachers try to bring students to the achievements of the program by conducting their lessons. At this point, it is very important to analyze the program acquisitions by considering some criteria. For this purpose, it is seen that chemistry education researchers have been examining chemistry curricula according to different criteria for many years (Aydın, 2013; Aydın et al., 2019; Ayyıldız et al., 2019; Demir et al., 2013; Nakiboğlu, 2021; Tüzün et al., 2019; Yaralı, 2022; Zorluoğlu et al., 2016). Ayyıldız et al. (2019) analyzed the acquisitions in the 2018 Chemistry Curriculum according to the Original and Revised Bloom Taxonomies. They determined that while the number of gains in the curriculum is quite high in the cognitive domain, the number of gains in the analysis, application, and knowledge level is low. It has been found that the number of gains in the synthesis and evaluation step is quite low. According to the cognitive process dimensions of the acquisitions, it was found that the acquisitions were mostly at the stage of understanding and analyzing, but there were very few gains at the stage of applying, remembering, and evaluating. In addition, it was found that there was very little gain in the creating step. Zorluoğlu et al. (2016) have also reached similar findings as a result of the analysis and evaluation of the 2013 Chemistry Curriculum acquisitions according to the structured Bloom taxonomy, and they have determined that was not acquisition for the creation step. Tüzün et al. (2019) examined the goals of the 2018 Chemistry Curriculum based on Toulmin argument pattern components and they used five argument pattern components as evaluating criteria. They found that each of the acquisitions of the high school chemistry teaching program consisted of at least one of Toulmin's argument pattern components and the construction of the acquisitions as arguments was average. Yaralı (2022) analyzed and compared the 2013 and 2018 secondary education chemistry course curricula in terms of the essential elements of the program which are “general objectives”, “units”, “recommended topic titles”, “numbers of learning outcomes”, “course hours”, “learning-teaching processes”, and “measurement and evaluation” sections.

It is seen that 13 general goals are included in the part related to the implementation of the 2018 Chemistry Curriculum. When these goals are examined, it is seen that the 9th item is as follows: "It is aimed for students to obtain data by experimenting, to make inferences using these data, to interpret them and to reach generalizations" (MoNE, 2018a). This shows that the 2018 Chemistry Curriculum has a program philosophy that constructs the knowledge rather than taking the knowledge readily, by making inferences through experimental studies, interpreting the data, and accessing theoretical knowledge in this way. In addition, when the questions in the PISA exam, which is one of the international exams, are examined, it is seen that an important part of the questions here are those that need to be inferred using scientific data, and answers should be produced accordingly (OECD, 2019).

Demir (2021) stated that the 2018 Secondary School Chemistry Curriculum and the 2018 Secondary Science High School Chemistry Curriculum were prepared in the same direction and that creating a solid chemistry infrastructure is dominant in both programs. She also drew attention to the fact that one of the main differences between the two programs is that more experimental applications are included in the Science High School Chemistry Curriculum compared to the Chemistry Curriculum. Of course, it is extremely important to have experimental studies in terms of the establishment purposes of science high schools. On the other hand, this difference should not be too much since experimental studies have an essential role in students' acquiring some 21st century skills. At the same time, as explained above, students are required to participate in experimental studies in order to achieve the philosophy and general objectives of the 2018 Chemistry Curriculum.

It is clear that the acquisitions of the curriculum must overlap with this goal and the philosophy of the Chemistry Curriculum, and accordingly, secondary school chemistry textbooks prepared in line with the acquisitions of the curriculum must be written in the same direction. When the secondary school chemistry textbooks are examined, it is seen that in order to get the approval of the Board of Education, the acquisitions of the chemistry curriculum are taken into account, and if an expression such as "experiment is done" is included in the acquisition, an experiment for the relevant acquisition is added to the textbooks. However, it is not known whether the general goals of the programs are taken into account too much. In particular, the explanations above are clear about how guiding the textbooks are for teachers. For this reason, chemistry teachers will take textbooks as a guide when they want to apply experiments in their classes. In this case, the other important point, as well as the inclusion of an experiment suitable for the learning outcome in the textbooks, is the overlap between the goals of the curriculum regarding laboratory studies and the experiments included in the textbooks.

With this in mind, firstly this study is aimed to examine to what extent the goal of the 2018 Chemistry Curriculum for experimental studies overlaps with the acquisitions of the curriculum and to what extent secondary school chemistry textbooks are prepared to meet this purpose. Secondly, it is also intended to compare this situation with the 2018 Secondary Science High School Chemistry Curriculum and determine the similarities and differences. The research problems that guide the study in line with these two aims are as follow.

1. What are the acquisitions concerning experimental studies in the 2018 Chemistry Curriculum and their distribution according to grades? Are there similarities and differences between the distribution of the 2018 Chemistry and the Science High School Chemistry Curricula?

2. What is the distribution of the experiments in the chemistry textbooks according to units and grades? Are there similarities and differences between the distribution of the chemistry textbooks and the science high school chemistry textbooks according to the units and grade levels of the experiments?

3. To what extent do the experiments in the chemistry textbooks meet the dimensions of 'gaining data by experimenting', 'inferring from the data, 'interpreting', and 'generalizing' in the 2018 Chemistry Curriculum's goal?

Method

The study's research design, participants, and the path followed in data collection and analysis are explained below.

Research Design

The study was designed according to the document analysis method. Bowen (2009) stated that document analysis is a systematic procedure for reviewing or evaluating printed and electronic materials. In addition, He pointed out that as with other analytical methods in qualitative research, document analysis requires the examination and interpretation of data in order to reveal meaning, gain understanding, and develop empirical knowledge.

Documents

In the study, the 2018 Chemistry Curriculum and the 9^{th}, 10^{th}, $11^{th,}$ and 12^{th} grade chemistry textbooks of the MoNE, written in line with this program and approved by the Board of Education, are the primary sources used for document analysis.

In the study, sampling was carried out according to criterion sampling, which is one of the purposive sampling methods. Criteria sampling of units handle the specified criteria (Büyüköztürk et al., 2008) and the criterion can be developed by the researcher (Yıldırım & Şimşek, 2018). Since a comparison will be made between grade levels, it was taken as a criterion for the selection of books written by the same publishing house, and therefore, 9^{th}, 10^{th}, $11^{th,}$ and 12^{th} grade chemistry textbooks published by the MoNE were included in the analysis. The list of analyzed books is given in Appendix 1 under the title of the book in the study.

Data Analysis

Altheide (1996), stated that the two interrelated principles that guide the document analysis process are impartiality and reliability, and the document analysis process is listed as follows: Determining the criteria to be included in the documents, collecting documents and data, determining the main analysis areas, coding, validating and analyzing the document (cited in Kıral, 2020). Two different analyzes were carried out in the study. The first of these is the analysis of the 2018 Chemistry Curriculum and the second is the analysis of secondary school chemistry textbooks. Separate paths were followed for both analyses and the results were compared to answer the research questions. The paths followed in both analyses are briefly explained below.

The path followed in the analysis of the program

First of all, the 2018 Chemistry Curriculum was analyzed for the acquisitions and explanations of the program in order to determine the number of acquisitions related to laboratory / experimental studies and their distribution according to grades. For this purpose, the analysis was carried out by accepting the acquisitions and explanations in the program as an acquisition related to laboratory/experimental studies, depending on the fact that they contain a statement about conducting experiments. In this way, after the analysis for each grade level, the findings were presented in tables according to the total number of acquisitions in the chemistry curriculum and the distribution of the laboratory acquisitions by grade level. Finally, the findings were compared with the analysis findings for the 2018 Science High School Chemistry Curriculum acquisitions.

The path followed in the analysis of textbooks

Textbook analyzes were carried out in two stages. In the first stage, the distribution and objectives of the experiments in the textbooks of each grade level were determined, and then each experiment was associated with acquisitions of the 2018 Chemistry Curriculum and tabulated separately for each grade level. Afterward, each grade level finding was compared with the findings obtained from science high school textbooks.

In the second stage of the textbook analysis, the analysis was carried out according to the statement "*Getting data by doing experiments, making inferences using these data, interpreting them and reaching generalizations*" in the 9th item of the 2018 Chemistry

Curriculum's aims. During this analysis, the trial and the rubric developed by Nakiboğlu (2021) were used. Explanations on this matter can be found in the related article (Nakiboğlu, 2021), and the rubric used is shown in Table 1.

Tablo 1. The rubric of Dimension Analysis

Dimensions in the 9th item of the program aim	Dimension analysis criterion
Obtaining data by experimenting	Basic BSB: observation, measurement, classification, data recording, number-space relationship, communication. Experiment validation BSB: prediction, variable identification, operational identification
Making inferences using data	Finding a result using raw data.
Interpretation	Reaching new relationships by combining the found result with theoretical knowledge.
Generalization	Making generalizations using the relationships reached.

Validity and reliability

To obtain intra-coder reliability, the author made all analyzes twice at different times and the differences were corrected. In addition, the chemistry curriculum analysis was repeated after the textbook analysis, and the related acquisitions were analyzed for the third time and it was determined that the results were completely compatible. Finally, encoder reliability was provided by making textbook analyses both in electronic format of textbooks and on printed textbooks (Gay & Airasion, 2000, s.175).

Findings

The findings obtained from the study are given below to answer the research questions.

Findings Regarding the First Research Question

In the first research question, the answer to the number of acquisitions of the 2018 Chemistry Curriculum for experimental studies and their distribution according to grades was sought. Findings related to this are presented in Table 2.

Table 2. The Total Number of Acquisitions of the 2018 Chemistry Curriculum and the Distribution of Experiment-related Acquisitions by Grade Levels

Grade level	Total number of acquisitions	Number of acquisitions related to the experimental study	%
9th grade	38	2	5,3
10th grade	23	8	34,8
11th grade	35	3	8,6
12th grade	31	4	12,9
Total	107	17	15,9

When Table 2 is examined, it is seen that the total number of acquisitions in the 2018 Chemistry Curriculum is 107 and 17 of these acquisitions, 15.9% of all program acquisitions, are related to experimental studies. When the percentages of acquisitions that include experimental studies at the grade level are examined, it is seen that 5.3% of the acquisitions in the 9th grade and 34.8% of the acquisitions in the 10th grade are related to experimental studies. It is understood that 40.1% and 28.1% of 11th and 12th grade acquisitions, respectively, are acquisitions involving experimental studies. The findings for comparing the experimental study-related acquisition percentages in the 2018 Chemistry Curriculum with the experimental study-related acquisition percentages in the 2018 Science High School Chemistry Curriculum are presented in Figure 1.

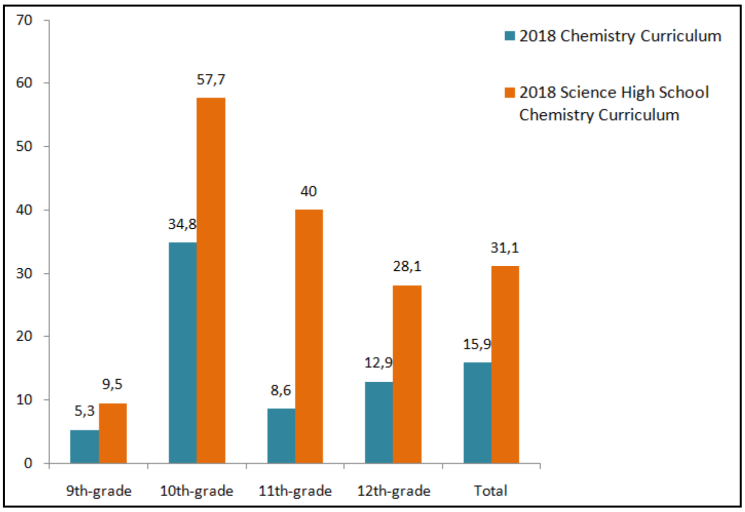


Figure 1. Comparison of the experimental study-related acquisition percentages in the 2018 Chemistry Curricula

When the two curricula are compared, it is seen that 5.3% of the achievements in the 2018 Chemistry Curriculum for 9th grade are related to experimental studies, while 9.5% of them are related to science high school. As can be seen from Figure 1, for both curricula, the least experimental work acquisition is in the 9th grade, while the most experimental work acquisition is in the 10th grade.

Findings Regarding the Second Research Question

In the second research question, it was investigated how the experiments in the 9th, 10th, 11th, and 12th grade chemistry textbooks were distributed according to the units and grade levels. The results of the analysis of the textbooks made for this purpose are presented in separate tables for each grade level (Tables 3, 4, 5, and 6). After the presentation of tables for each grade level, the findings were compared with the findings of the 2018 Science High School Chemistry Curriculum. The distribution of the experiments in the 9th grade chemistry textbook by units and findings of the related acquisitions are shown in Table 3.

Tablo 3. The Distribution of the Experiments in the 9th grade Chemistry Textbook by Units and Findings of the Related Acquisitions

Experiment No	Unit	Title of Experiment	Associated Acquisition
1	States of matter	Effect of temperature on viscosity	9.4.3.2. To explains the factors affecting viscosity in liquid. c. *Viscosity experiments of water, glycerin, and olive oil at different temperatures are carried out and the results obtained are compared.*
2		The phase change of pure water	9.4.4.3. To interprets the state change graphs of pure substances. *c. The state change experiment of pure water is done and the graph is drawn.*

The grade 9 chemistry textbook contains five units. When Table 3 is examined, there are 2 experiments in total in the 9th grade chemistry textbook, and these experiments are placed in the 4th unit, "States of matter". It is seen that there are no experiments in the units of “Chemistry science”, “Atomic and periodic system”, “Chemical interactions between species” and “Nature and chemistry” in the 9th grade chemistry textbook. Looking at the related acquisitions of these experiments in the Chemistry Curriculum from Table 3, it is seen that there are only 2 acquisitions related to the experimental studies for the “States of matter" unit.

When these findings are compared with the achievements of the 9th grade Science High School Chemistry Curriculum and the experiments in the textbook, it is seen that while there are four outcomes for science high school, there are five experiments in the textbook. The common acquisitions and experiments are the same for the two programs and acquisitions, and it has been determined that there are three extra experiments in the 9th grade science high school textbook. The distribution of the experiments in the 10^{th} grade chemistry textbook by units and findings of the related acquisitions are shown in Table 4.

Table 4. The Distribution of the Experiments in the 10^{th} grade Chemistry Textbook by Units and Findings of the Related Acquisitions

No	Unit	Title of experiment	Associated acquisition
1	Fundamental laws of chemistry and chemical calculations	Preparing of iron (II) sulfide	10.1.1.1. To explain the fundamental laws of chemistry. *b. The experiment concerning obtaining the iron(II) sulfide compound is done.*
2		Precipitation of lead (II) iodide	10.1.3.1. To explain chemical reactions. *c. The experiment concerning the precipitation of lead(II) iodide is done.*
3	Mixtures	Dissolution of different substances in water	10.2.1.2. To explain the dissolution process at the molecular level. c. *The experiments concerning the dissolution of different substances (sodium chloride, ethyl alcohol, carbon tetrachloride) in water are done.*
4		Separation of the iron powder-sulfur powder mixture	10.2.2.1. To explain mixture separation techniques used in industry and health fields.
5		Separation by melting point difference	b. *Experiments to separate the mixtures are done.*
6		Separation using the particle size difference	
7		Separation using the boiling point difference	
8		Separation using the solubility difference	
9		Separation using the solubility difference	
10		Separation by density difference	
11	Acids, Bases, and Salts	Effect of acids and bases on colored substances	10.3.1.1. To distinguish acids and bases with the help of their known properties. *c. Indicator concept and pH paper are introduced by doing experiments that acids and bases change the color of some colored substances (tea, grape juice, red cabbage).*
12		Using pH paper	*ç. The acidity or alkalinity values of vinegar, lemon juice, bleach, sodium hydroxide, hydrochloric acid, and sodium chloride solutions are interpreted using pH paper.*
13		Salt formation	10.3.2.1. To explain the reactions between acids and bases. *b. The concepts of acid, base, and salt are related by doing an experiment for the formation of sodium sulfate from the interaction of sodium hydroxide and sulfuric acid.*
14		Amphoteric property of aluminum metal	10.3.2.2. To explain the important reactions of acids and bases in terms of daily life. *b. An experiment is done showing the amphoteric property of the aluminum metal.*
15		Effect of lime and caustic on oil, hair, and skin	10.3.3.1. To explain the benefits and harms of acids and bases. *b. The effect of lime and caustic on oil, hair and skin is explained by experimentation.*

10th grade chemistry textbook contains four units. From Table 4, it is seen that there are eight acquisitions in the three units of the 2018 Chemistry Curriculum and 15 experiments related to these acquisitions in the 10th grade textbook. Experiments are located in the other three units except for the last unit, “Chemistry is everywhere”. In the first unit, 'Fundamental laws of chemistry and chemical calculations', there are two acquisitions and two experiments related to these acquisitions. The second unit is the 'Mixtures' unit and there are two acquisitions and eight experiments related to these acquisitions. The number of experiments in the third unit, ‘Acids, bases, and salts’, is five and the number of related acquisitions is four.

When the findings in Table 4 are compared with the number of experiments and related acquisitions in the 10th grade science high school chemistry textbook, there are 15 acquisitions and 22 experiments related to these acquisitions in 10th grade science high school. While there are experiments for three units in chemistry textbooks, it is seen that there are experiments for all units, namely four units, in the science high school chemistry textbook. The distribution of the experiments in the 11th grade chemistry textbook by units and findings of the related acquisitions are shown in Table 5.

Table 5. The Distribution of the Experiments in the 11th grade Chemistry Textbook by Units and Findings of the Related Acquisitions

No	Unit	Title of experiment	Associated acquisition
1	Gases	Diffusion of gases	11.2.3.1. To explain gas behavior via kinetic theory. *c. Diffusion experiment is done; …*
2	Liquid Solutions and Solubility	Comparison of boiling points of pure water and solutions of different concentrations	11.3.3.1. To establishes a relationship between the colligative properties of solutions and their concentrations. *c. Boiling point determination experiments of pure water and aqueous solutions of different concentrations are done.*
13	Equilibrium in Chemical Reactions	Strong Acid-base titration	11.6.3.8. To determines strong acid/base concentrations by titration method. *a. A titration experiment is performed and the results are interpreted by showing them graphically.*

The 11th grade textbook consists of six units, and as can be seen from Table 5, there are three experiments for only three units in the chemistry textbook. The units in which the experiments take place are units of "Gases", "Liquid solutions and solubility" and "Equilibrium in chemical reactions". It is seen that the total number of acquisitions related to the experiments in the chemistry textbook is three. It has been determined that there is no experiment in the first unit, “Modern atomic theory”, in the 4th unit, in the "Energy in chemical reactions" unit, and in the 5th unit, in the "Rate in chemical reactions" unit. When the findings of Table 5 are compared with those of the science high school, there are 14 acquisitions in the 2018 Science High School Curriculum for the 11th grade and 17 experiments related to these acquisitions in the science high school chemistry textbook. The distribution of the experiments in the 12th grade chemistry textbook by units and findings of the related acquisitions are shown in Table 6.

Table 6. The Distribution of the Experiments in the 12th grade Chemistry Textbook by Units and Findings of the Related Acquisitions

No	Unit	Title of experiment	Associated acquisition
1	Chemistry and Electricity	Electrochemical battery experiment	12.1.2.1. To explain the concepts of the electrode and electrochemical cell. *d. Zn/Cu electrochemical cell experiment is done;…..*
2		Copper plating by electrolysis method	12.1.5.1. To explain electrolysis in terms of electric current, time, and mass of matter undergoing change. *d. Plating experiment is done.*
3		Electrolysis of water	12.1.5.2. To explain the process of obtaining chemical substances by electrolysis method. *The experiment of hydrogen and oxygen production via electrolysis of water is done.*
4	Organic Compounds	To make the soap	12.3.7.1. To explain the names, formulas, and usage areas of esters. *d. The experiment of soap making is done.*

There are four units in total in the 12th grade. As can be seen from Table 6, a total of four experiments were included in only two units. It is seen that there are three experiments in the first unit, 'Chemistry and electricity', and there are three achievements related to these experiments. One experiment is included in the textbook for one achievement in the 3rd unit, 'Organic compounds'. It was determined that in the second unit, 'Introduction to carbon chemistry' and the last unit, 'Energy resources and scientific developments', there was no experiment in the 12th grade upper-secondary school textbook and related achievements involving any experiment in the 2018 Chemistry Curriculum. When these findings are compared with those of the science high school, it is seen that there are four units in the 12th grade science high school textbook and a total of 10 experiments in three of these units. In the first unit, 'Chemistry and electricity', there are six experiments and six achievements associated with these experiments. In the second unit, "Introduction to carbon chemistry", there are two experiments corresponding to a single outcome. Two experiments are included in the 12th grade science textbook for two achievements in the 3rd unit, 'Organic compounds' of the 2018 Science High School Curriculum.

The findings for the comparison of the 9th grade science high school chemistry textbook and the 9th grade chemistry textbook in terms of the number of experiments and their distribution to the units are given in Figure 2.

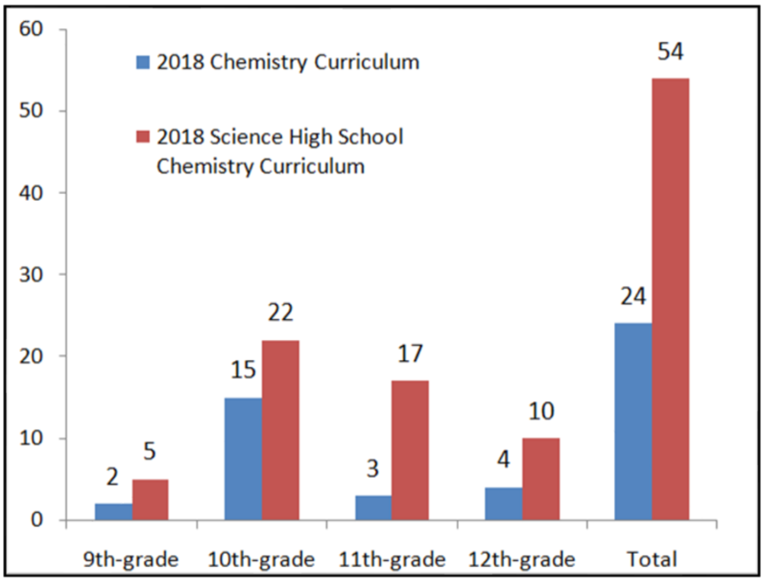


Figure 2. Comparison of 9th grade science high school chemistry textbook and 9th grade chemistry textbook in terms of distribution and number of experiments according to grades

When the science high school chemistry textbook and the chemistry textbook are compared in terms of the number of experiments in the textbooks, it was found that while there are 54 experiments in the science high school chemistry textbook, there are 24 experiments in the chemistry textbook.

Findings Regarding the Third Research Question

In the third research question, it was investigated to what extent the experiments/activities in the chemistry textbooks met the dimensions of "gathering data by experimenting", "making inferences using the data", "interpreting" and "generalizing". For this purpose, firstly, analyses of the textbooks for each class are shown in separate tables below (Tables 7, 8, 9, and 10) and then, the findings of the experiments in all classes regarding these dimensions were compared and the findings are shown in Table 11.

Table 7. Findings Regarding Dimensional Analysis of Experiments/Activities in the 9th grade Chemistry Textbook

Experiment No	Basic SPS						Experiment confirmation SPS			Make inferences	Interpretation	Generalization
	O	M	Cl	RD	N/S	Co	P	VI	OD			
1	-	+	-	+	-	-	-	-	-	+	+	+
2	+**	+	-	+	-	+*	-	-	-	+	-	-
Total	1	2	-	2	-	1	-	-	-	2	1	1

* Drawing graphics
**Not direct observation

As can be seen from Table 7, in the 9th grade chemistry textbook, students are expected to make observations only in one experiment, and this observation was not directly expressed as "observe", and it was decided by inferring from the explanations that this is an observation skill (O). It was determined that they were expected to measure (M) and record data (RD) in two experiments and in one experiment there was the procedure for communication skill (Co). On the other hand, it can be seen from Table 7 that no guidance has been made for the development of basic SPS such as classification (Cl) and number-space relationship (N/S), and also experimental confirmation SPS such as prediction (P), variable determination (VI) and operational definition (OD). It was determined that in two of the experiments, students were asked to make inferences, and in one experiment, they were asked to interpret and generalize the result of the statement.

Table 8. Findings Regarding Dimensional Analysis of Experiments/Activities in the 10th grade Chemistry Textbook

Experiment No	Basic SPS						Experiment confirmation SPS			Make inferences	Interpretation	Generalization
	O	M	Cl	RD	N/S	Co	P	VI	OD			
1	+	+	-	+	+	+	-	-	-	+	-	+
2	+	+	-	+	-	-	-	-	-	+	+	-
3	+	+	-	+	-	-	-	-	+	+	+	-
4	+	-	-	-	-	-	-	-	-	+	-	+
5	-	+	-	-	-	-	-	-	-	-	-	-
6	-	+	-	-	-	-	-	-	-	+	-	+
7	+	+	-	+	-	+*	-	-	-	-	+	-
8	+	+	-	-	-	-	-	-	-	+	-	-
9	-	+	-	-	-	-	-	-	-	+	-	-
10	+	+	-	-	-	-	-	-	-	+	-	-
11	+	+	-	+	-	+	-	-	-	+	+	+
12		+	+	+	-	+	-	-	-	+	+	+
13	+	+	-	-	+		-	-	-	+	-	-
14	+	+	-	-	-	-	-	-	-	+	-	-
15	+	-	-	-	-	-	-	-	-	+	+	-
Total	11	13	1	6	2	4	-	-	1	13	6	5

As can be seen from Table 8, in 11 of the 15 experiments in the 10th grade chemistry textbook, students were required to observe (O), measure in 13 experiments, classify (Cl) in one experiment, record data (RD) in six experiments, make numerical calculations (N/S) in two experiments, and a process that would improve communication skill (Co) in four experiments are expected to do. On the other hand, while no guidance was given to the development of experimental confirmation PPS, such as prediction (P) and variable identification (VI) skills in experiments, operational definition (OD) was included in the experiment. While students were expected to make inferences in 13 of the experiments, it was determined that they were asked to interpret the results of the experiments in six experiments and to generalize the results in five experiments.

Table 9. Findings Regarding Dimensional Analysis of Experiments/Activities in the 11th grade Chemistry Textbook

Experiment No	Basic SPS						Experiment confirmation SPS			Make inferences	Interpretation	Generalization
	O	M	Cl	RD	N/S	Co	P	VI	OD			
1	+	+	-	+	+		-	-	-	+	+	+
2	-	+	-	+	-	-	-	-	-	+	+	+
3	-	+	-	+	+	+*	-	-	-	+	+	-
Total	1	3	-	3	2	1	-	-	-	3	3	2

* Drawing graphics

As seen in Table 9, students are expected to make observation (O) in only one of three experiments in the 11th grade chemistry textbook, to measure (M) in three experiments, to record data (RD) in three experiments, to make numerical calculations (N/S) in two experiments, and to interpret graphics in a way to improve communication skill (Co) in one experiment. In all of the experiments, there is no guidance for the classification skill (Cl) from the basic BSB. In addition, it is seen that no direction was made for experiment confirmation SPS development in any of the experiments. While students were expected to make inferences and interpretations in all of the experiments, it was determined that they were asked to generalize the findings of the experiment in two experiments.

Table 10. Findings Regarding Dimensional Analysis of Experiments/Activities in the 12th grade Chemistry Textbook

Experiment No	Basic SPS						Experiment confirmation SPS			Make inferences	Interpretation	Generalization
	O	M	Cl	RD	N/S	Co	P	VI	OD			
1	+	+	-	-	-	+	-	-	-	+	+	+
2	+	+	-	-	-	-	-	-	-	+	+	+
3	-	+	-	-	+	-	-	-	-	+	-	-
4	-	+	-	-	-	-	-	-	+	+	-	-
Total	2	4	-	-	1	1	-	-	1	4	2	2

As can be seen from Table 10, in two of the four experiments in the 12th grade chemistry textbook, students are expected to make an observation (O), make a measurement (M) in all experiments, make numerical calculation (N/S) in one experiment, and perform operations that will improve communication skill (Co) in one experiment. In all of the experiments, there is no guidance from the basic BSB to the ability to classify (Cl) and data recording (RD) skills. In addition, it can be seen from Table 10 that none of the experiments provided any guidance for the development of experiment confirmation SPS such as prediction (P), and variable identification (VI) skills. In only one experiment, students are expected to perform operations related to operational definition (OD) skill. While the students were expected to make inferences in all of the experiments, it was determined that they were asked to interpret the results of the experiment in two experiments and to generalize the results in two experiments.

Finally, the comparison of the results of the dimensional analysis of the experiments in the textbooks at all grade levels is given in Table 11.

Table 11. Comparison of Experiments/Activities in Chemistry Textbooks according to Dimensional Analysis

Grade Level	Number of experiments	Basic SPS						Experiment confirmation SPS			Make inference	Interpretation	Generalization
		O	M	Cl	RD	N/S	Co	P	VI	OD			
9	2	1	2	-	2	-	1	-	-	-	2	1	1
10	15	11	13	1	6	2	4	-	-	1	13	6	5
11	3	1	3	-	3	2	1	-	-	-	3	3	2
12	4	2	4	-	-	1	1	-	-	1	4	2	2
Total	24	15	22	1	11	5	7	-	-	2	22	12	10

As can be seen from Table 11, as a result of the Basic SPS analysis included in the first dimension, the most skill type was 'measurement' (M) and in 22 experiments the measuring skill was addressed, followed by 'observation' (O), 'data recording' skill (RD). followed by communication (Co) and the 'number/space relation' (N/S). It is seen that there is only one experiment for which action is requested for the 'Classification' (Cl) skill. It was determined that experiment confirmation SPS was only asked to make an 'operational definition ' (OD). In none of the experiments, it was seen that no guidance was given for predicting (P) and variable identification (VI) skills.

Conclusions and Suggestions

In the study, firstly, the acquisitions of the 2018 Chemistry Curriculum including laboratory/experimental studies were examined, and it was concluded that the acquisitions that include explanations about experimental studies are approximately 16% of the total acquisitions of the curriculum. When the percentages of the acquisitions related to the experimental studies are examined according to the grade levels, it was determined that the grade level with the most experimental work in the explanation part was the 10th grade, and this percentage was approximately 35%. This situation can be considered as an indication that the 2018 Chemistry Curriculum partially attaches importance to experimental studies. When this result is compared with the percentage of the acquisitions in the 2018 Science High School Curriculum related to the experimental study, it is seen that the Science High School Curriculum's acquisitions are twice (31.1%) compared to the 2018 Chemistry Curriculum's acquisitions (Nakiboğlu, 2021). The fact that the experimental study acquisitions are high in Science High School Curriculum is compatible with the established purpose of the science high school and the philosophy of this curriculum. It has been also determined that the number of acquisitions related to experimental studies in the 2018 Science High School Curriculum at each grade level is higher than the 2018 Chemistry Curriculum, the number of acquisitions involving the least experimental studies in both curricula is at the 9th grade level, and the number of acquisitions containing the most experimental studies is at the 10th grade level. The importance of experimental studies for science high school students is revealed with the following statement under the title of the basic philosophy and general objectives of the 2018 Science High School Chemistry Curriculum.

"…With the Science High School Chemistry Curriculum, which was prepared in accordance with the emerging needs, it is aimed that students spend more time in the laboratory environment and prepare projects to encourage them to become scientists and do scientific studies" (MoNE, 2018b, p.12).

Although students in science high schools need to spend more time in the laboratory, the important point for other high schools is how the experiments in this process are carried out. In acquiring scientific knowledge, students should not only take part in laboratory studies but also these studies should be carried out based on research and inquiry. This is important for students attending other types of high school as well as for students attending science high school. For this reason, it is extremely important to provide guidance and to plan in a way that will provide many skills in the conduct of the experiments. For this reason, the preparation of experimental studies in school textbooks as a guide for teachers and their use will be the most important basis for structuring scientific knowledge. Based on these considerations, in the second part of this study, dimensional analyzes for experiments in chemistry textbooks are also included.

The experiments placed in the chemistry textbooks were analyzed for the four dimensions "gathering data by experimenting", "making inferences using the data", "interpreting" and "generalizing". The first dimension, "gathering data by experimenting", is directly related to the basic SPS and experiment confirmation SPS. As a result of this analysis for SPSs, it is seen that the experiments mostly focus on the basic SSBs, and they focus especially on “observation” and “measurement” skills. Experiments do not seem to give much space to high-level SPS development. Basic SPSs form the basis of high-level skills (Padilla, 1990) and it has been stated that these skills can be acquired by students starting from the pre-school period, while high-level skills can be gained from the second level of primary education (Aydoğdu et al., 2012). For this reason, it can be said that it is much more important for the experiments in chemistry textbooks to focus on higher SPSs. On the other hand, In none of the experiments, it was concluded that there was no guidance for improvement of experiment confirmation SPS such as predicting and variable identification skills. So, it can be said that chemistry textbooks do not give the necessary importance to the development of SPS. Padilla et al. (1984) also emphasized that scientific process skills should be strongly included in science programs and classrooms, whether they are considered as the way scientists think or as survival strategies for a changing world.

Another important conclusion about the study belongs to the textbook analysis of the other three dimensions. It was determined that inferences were made using the data in 22 of the 24 experiments in the chemistry books, interpretations were made in 12 experiments, and the results were generalized in only 10 experiments. Based on these results, at the end of the experiments in the chemistry textbooks, although students are expected to make inferences based on the results of the experiments, it is not required to interpret and generalize the results in half of the experiments. Özmen (2004), in his study on the reasons why teachers do not use the laboratory in chemistry classes, determined that the instructions and explanations about the experiments in the books were not found sufficient for nearly half of the teachers. In addition, it was determined that the experiments in the textbooks did not guide the students much and the teachers thought that these experiments were not based on research and inquiry. Looking at the years of this study, it can be said that although the programs have changed in the intervening 18 years, the guidance of the chemistry textbooks regarding the experiments has not changed much.

At the end of the study, it can be said that the 2018 Chemistry Curriculum partially meets the purpose of including experimental studies in both its purpose and the explanations in the acquisitions. Although there is a quantitative harmony between the books and the outcome, when the experimental contents are examined, it can be said that this harmony does not fully comply with the curriculum's goal, and the way the experiments are given in the chemistry textbooks is not written in a way to fully guide the teachers. For this reason, the first suggestion is for the parts written about experimental studies during the writing of the textbooks.

In the writing of experimental activities in chemistry textbooks, a more skill development-oriented teacher and teacher-guiding way should be adopted. In fact, as it was seen in the previous study on science high schools (Nakiboğlu, 2021), it does not seem possible to carry out this number of experiments during the course hours in the program and to follow a path based on research and inquiry. For this reason, to achieve this goal of the program, it can be suggested that separate course hours should be included in the program, and guidebooks should be prepared to guide teachers.

Kimya Ders Kitaplarındaki Deneylerin 2018 Kimya Dersi Öğretim Programının Deneysel Çalışmaya Yönelik Amacı Açısından Analizi ve Fen Lisesi Durumu ile Karşılaştırılması

Özet:

Bu çalışma, 2018 Kimya Dersi Öğretim Programının deneysel çalışmalar hedefinin programın kazanımlarıyla ne ölçüde örtüştüğünü ve lise kimya ders kitaplarının bu amaca yönelik olarak ne ölçüde hazırlandığını incelemeyi amaçlamaktadır. Çalışma aynı zamanda kimya ders kitaplarındaki etkinliklerin 2018 Kimya Dersi Öğretim Programının dört boyutunu (deneyerek veri toplama, verileri kullanarak çıkarım yapma, yorumlama ve genelleme)) ne ölçüde karşıladığını değerlendirmeyi de amaçlamaktadır. Son olarak bu durumun 2018 Fen Lisesi Kimya Dersi Öğretim Programı ile karşılaştırılması, benzerlik ve farklılıklarının belirlenmesi hedeflenmiştir. Bu amaçlar doğrultusunda öncelikle 2018 Kimya Dersi Öğretim Programında deneysel çalışmaları içeren kazanımlar belirlenmiştir. Daha sonra 9., 10., 11. ve 12. sınıf kimya ders kitaplarındaki deneyler analiz edilmiştir. Ders kitabı analizi için iki yol kullanılmıştır. İlk aşamada 2018 Kimya Dersi Öğretim Programındaki kazanımlar ile ders kitaplarındaki deneylerin ne ölçüde örtüştüğü ve deneylerin ünitelere göre nasıl dağıldığı belirlenmiştir. İkinci aşamada, ders kitaplarında yer alan deneyler, 2018 Kimya Dersi Öğretim Programının dört boyutuna göre bir dereceli puanlama anahtarı aracılığıyla analiz edilmiştir. Çalışmada, deneysel çalışmalarla ilgili açıklamalar içeren kazanımların 2018 Kimya Dersi Öğretim Programı toplam kazanımlarının yaklaşık %16'sına karşılık geldiği sonucuna ulaşılmıştır. Ayrıca, BSB analizi sonucunda, deneylerin daha çok temel BSB'lere odaklandığı, özellikle "gözlem" ve "ölçüm" becerilerine odaklanıldığı belirlenmiştir.

Anahtar kelimeler: 2018 Kimya Dersi Öğretim Programı, lise kimya ders kitabı, deneysel çalışma.

References

Ayas, A. (1995). Fen bilimlerinde program geliştirme ve uygulama teknikleri üzerine bir çalışma: İki çağdaş yaklaşımın değerlendirilmesi. *Hacettepe Üniversitesi Eğitim Fakültesi Dergisi, 11*(11), 149-155.

Aydın, A. (2013). Representation of science process skills in the chemistry curricula for grades 10, 11 and 12/Turkey. *International journal of Education and Practice, 1*(5), 51-63.

Aydın, A., Ayyıldız, Y., & Nakiboğlu, C. (2019). Investigation of the gains of the 2018 Science High School Chemistry Curriculum according to the revised Bloom's Taxonomy and Comparison with 2018 Chemistry Curriculum. *Necatibey Faculty of Education Electronic Journal of Science and Mathematics Education, 13*(2), 1186-1215.

Aydoğdu, B., Tatar, N., Yıldız, E., & Buldur, S. (2012). İlköğretim öğrencilerine yönelik bilimsel süreç becerileri ölçeğinin geliştirilmesi. *Journal of Theoretical Educational Science, 5*(3), 292-311.

Ayyildiz, Y., Aydin, A., & Nakiboğlu, C. (2019). 2018 Yılı Ortaöğretim Kimya Dersi Öğretim Programı kazanımlarının orijinal ve yenilenmiş Bloom Taksonomisine göre incelenmesi. *Mehmet Akif Ersoy Üniversitesi Eğitim Fakültesi Dergisi*, 52, 340-376.

Bowen, G. A. (2009). Document analysis as a qualitative research method. Qualitative Research Journal, 9(2), 27-40.

Büyüköztürk, Ş., Kılıç Çakmak, E., Akgün, Ö. E., Karadeniz, Ş., & Demirel, F. (2008). Bilimsel araştırma yöntemleri. Pegem Akademi publishing.

Demir, E. (2021). 2018 Ortaöğretim Kimya Dersi Öğretim Programı ve 2018 Ortaöğretim Fen Lisesi Kimya Dersi Öğretim Programı'nın temel ögeler açısından karşılaştırılması. *Turkiye Kimya Dernegi Dergisi Kısım C: Kimya Egitimi*, *6*(2), 171-207. https://doi.org/10.37995/jotcsc.989550

Demir, E., & Nakiboğlu, C. (2021). 2018 Yılı Fen Bilimleri Dersi Öğretim Programı'nın kimya konuları bağlamında incelenmesi. Journal of Turkish Chemical Society Section: C, 6(1), 23-70.

Demir, E., Gacanoğlu, Ş., & Nakiboğlu, C. (2013). 2013 Kimya Dersi Öğretim Programı'na yönelik öğretmen görüşleri doğrultusunda 2017 Kimya Dersi Öğretim Programı'nın değerlendirilmesi. *Turkiye Kimya Dernegi Dergisi Kısım C: Kimya Egitimi, 2*(2), 135-184.

Gay, L. R. & Airasion, P. (2000). Educational research: Competencies for analysis and application. Upper Saddle River: Prentice-Hall.

Kıral, B. (2020). Nitel bir veri analizi yöntemi olarak doküman analizi. *Sosyal Bilimler Enstitüsü Dergisi, 15,* 170-189.

Mattheis, F. E., & Nakayama, G. (1988). Effects of a laboratory-centered inquiry program on laboratory skills, science process skills, and understanding of science knowledge in middle grades students. Reports – Research, ED307148

Ministry of National Education (MoNE) (2018a). Ortaöğretim Kimya Dersi (9, 10, 11 ve 12. sınıflar) Öğretim Programı.

Ministry of National Education (MoNE) (2018b). Fen Lisesi Kimya Dersi (9, 10, 11 ve 12. sınıflar) Öğretim Programı.

Nakiboğlu, C. (2021). Fen lisesi kimya ders kitaplarındaki deneylerin 2018 Yılı Fen Lisesi Kimya Dersi Öğretim Programı'nın deneysel çalışmaya yönelik amacı açısından analizi. *Turkiye Kimya Dernegi Dergisi Kısım C: Kimya Egitimi, 6*(2), 209-239. https://doi.org/10.37995/jotcsc.991061

Nakiboğlu, C., & Şen, A. Z. (2020). Kimya öğretmen adaylarının kimya laboratuvarlarında kullanılabilecek öğretim strateji ve modellerine yönelik görüşlerinin incelenmesi. *Necatibey Faculty of Education Electronic Journal of Science & Mathematics Education,* 14(1), 717-760. Doi: 10.17522/balikesirnef.746186

OECD (2019). *PISA 2018 Assessment and Analytical Framework*. PISA, OECD Publishing, Paris, https://doi.org/10.1787/b25efab8-en.

Özmen, H. (2004). Kimya-1, 2 ders kitaplarındaki deneylerin uygulanabilirlik düzeylerine ve laboratuvar kullanımına yönelik öğretmen görüĢleri. *Hasan Ali Yücel Eğilim Fakültesi Dergisi, 1,* 11-27.

Padilla, M. J., Okey, J. R., & Garrard, K. (1984). The effects of instruction on integrated science process skill achievement. *Journal of Research in Science Teaching, 21*(3), 277–287. https://doi.org/10.1002/tea.3660210305.

Padilla, M. J. (1990). The science process skills. *Research Matters to the science Teacher*. ERIC Document Reproduction Service No ED266961.

Tüzün, Ü. N., Bilir, V., & Eyceyurt, G. (2019). Ortaöğretim Kimya Dersi Öğretim Programı kazanımlarinin Toulmin Argüman Modeli bileşenlerine göre değerlendirilmesi. *Abant İzzet Baysal Üniversitesi Eğitim Fakültesi Dergisi*, *19*(4), 1322-1333.

Uzezi, J. G., & Zainab, S. (2017). Effectiveness of guided-inquiry laboratory experiments on senior secondary schools students academic achievement in volumetric analysis. *American Journal of Educational Research, 5*(7), 717-724.

Yaralı, D. (2022). Comparison of 2013 and 2018 Chemistry Course Curricula in terms of the basic elements of the program. Journal of Turkish Chemical Society Section C: Chemistry Education (JOTCS-C), 7(2), 153-182.

Yıldırım, A., & Şimşek, H. (2018). Sosyal bilimler nitel araştırma yöntemleri. Seçkin publishing.

Zorluoğlu, S. L., Kizilaslan, A., & Sözbilir, M. (2016). Ortaöğretim kimya dersi öğretim programı kazanımlarının yapılandırılmış Bloom taksonomisine göre analizi ve değerlendirilmesi. *Necatibey Eğitim Fakültesi Elektronik Fen ve Matematik Eğitimi Dergisi*, 10(1), 260-279.

Necatibey Eğitim Fakültesi Elektronik Fen ve Matematik Eğitimi Dergisi
Cilt 16, Sayı 2, Aralık 2022, sayfa 428-455. ISSN: 1307-6086
Necatibey Faculty of Education Electronic Journal of Science and Mathematics Education
Vol. 16, Issue 2, December 2022, pp. 428-455. ISSN: 1307-6086

Research Article

Prospective Mathematics and Physics Teachers' Experiences of Implementing Peer-Assessment

Olga GKIOKA

Boğaziçi Üniversitesi, North Campus, 34342 Bebek, Istanbul, Türkiye,
olga.gkioka@boun.edu.tr, http://orcid.org/0000-0002-0477-3217

Received : 07.06.2022 Accepted : 15.12.2022
Doi: 10.17522/balikesirnef.1127523

Abstract – The aim of the presented research is to investigate how prospective mathematics and physics teachers have experienced peer-assessment. The study took place within the context of one classroom assessment course designed for pre-service teacher education. Ninety-four (94) prospective teachers in four cohorts participated. They were giving written feedback to peer's coursework, writing reflective journals and participating in semi-structured interviews. The participants reported that grading is more difficult than giving feedback because grades are important as they "count" for students' success. Secondly, during the course, their thinking "shifted" to the idea that giving comments is more crucial than grading, because feedback helped them understand the assessment criteria and subsequently, improve coursework. Based on the findings, the main argument of the paper is about the role of pre-service teacher education for the development of classroom assessment skills and assessment knowledge in teachers.

Key words: assessment criteria, feedback, grading, peer-assessment, pre-service mathematics and physics teacher education.

Introduction

The presented study is heavily influenced by the work and research on classroom assessment by Brookhart and her colleagues (Brookhart, 2011; Brookhart et al., 2016) and on peer-assessment assessment by Black and Wiliam (1998), Brookhart (2017) and Shepard and her colleagues (2005). According to Brookhart and colleagues, through classroom assessment, the teacher should aim to enhance and support student learning. This study has used the term "classroom assessment" to include both formative and summative assessment. Brookhart and McMillan (2020) developed a broad definition of classroom assessment as follows:

Classroom assessment is a process that teachers and students use in collecting, evaluating, and using evidence of student learning for a variety of purposes, including diagnosing student strengths and weaknesses, monitoring student progress toward meeting desired levels of proficiency, assigning grades, providing feedback to students and parents, and enhancing student learning and motivation (Brookhart & McMillan, 2020, p. 4).

Assessment requires not only teachers' but also students' active involvement in assessment when they look at each other's work to identify strengths and weaknesses, to give feedback for improvement (peer-assessment) and then improve their own work (self-assessment). The primary goal is to promote learning and advance learners' responsibility for their own learning and development of metacognitive skills. Within the same framework, Brookhart (2017) has also developed the argument that grading and summative assessment should not mean marking only; grading should also guide and support learning. Hence, teachers need to develop grading practices which focus on learning and aim towards support of learning.

The literature review starts with one definition of peer-assessment and discusses how peer-assessment is related to assessment criteria, feedback and grading. Then, research on peer-assessment has been presented. The literature review concludes with research on teacher assessment practices. It is worth noting that the reviewed research studies were conducted in Turkey, as well as Europe and the US.

Peer-Assessment: Assessment Criteria, Feedback and Grading

During the process of peer-assessment, the intention is that both assessee and assessor to benefit from the process. In peer-assessment, students provide and receive feedback more promptly. Topping (2009) defined peer-assessment as follows:

Peer assessment is an arrangement for learners to consider and specify the level, value, or quality of a product or performance of other equal status learners. Products to be assessed can include writing, oral presentations, portfolios, test performance, or other skilled behaviors (p. 20).

Sadler (1989), in his seminal paper, discussed the three elements required for peer-assessment to promote learning:

• Firstly, students need to develop a clear view of the learning goals and a shared understanding of quality work similar to that of the teacher. In simple learner's words: "Where am I trying to go?"

• Secondly, information is needed about the present state of the learner and the expected one. The learner should be able to compare the current level of performance by using quality criteria. In student's simple terms: "Where am I?"

• Lastly, the learner should be able to take action to close the gap. "How best to get there?"

Sadler (1989) stressed the potential value of peer work both because the communication between students is in a language that students themselves naturally use, and also because students learn from taking the role of teachers and examiners. On such basis, Sadler (1989) strongly argued that self- and peer-assessment are of central importance to formative assessment. At the same time, Sadler argued that students must understand the criteria according to which their work will be judged. In a later paper, Sadler (1998) emphasized the need for the feedback to be understood by the learner and that students should be actively involved in further improvement and learning.

Along the same line, Boud (2000) argued that

Unless students are able to use the feedback to produce improved work, through for example, redoing the same assignment, neither they, nor those giving the feedback will know it has been effective (p. 158).

Feedback is only successful if students use it to improve their performance. As such, feedback needs to identify the quality of the work required and then, it needs to indicate the gap between the current achievement (i.e., strengths) and the desired performance (in relation to assessment criteria and the expected quality). Having discussed the important role of peer-assessment in the service of learning, the focus of the discussion will move to research on peer-assessment.

Research on Peer-Assessment

Research studies on peer-assessment have shown the value and the importance of student involvement in the process. They have provided strong evidence that peer-assessment and good quality feedback can promote learning and encourage progress in student learning in secondary education (i.e., Black & Harrison, 2004; Hodgen & Wiliam, 2006; Ketonen et al., 2020) and in higher education (Sluijsmans & Prins, 2006; Sluijsmans et al., 2002). Wiliam and Leahy (2015) labelled peer-assessment with the expression "activating students as learning sources for one another" and self-assessment with "activating students as owners of their own learning". Peer-assessment helps learners to carry on with the "next" step in the route of their learning, which is self-assessment. Yet, peer-assessment may help students develop ownership of learning when they are actively involved in assessing peer's work (Shepard, 2000).

According to research studies on formative assessment, feedback is an essential component of assessment because it has major influence on learning (Black, 2003; Black & Wiliam, 1998; Hattie, 2011). Shute (2008) used the term "formative feedback" and defined it as

information communicated to the learner that is intended to modify his or her thinking or behaviour for the purpose of improving learning (p. 154).

Butler (1988) showed that students get distracted by grades that go with written comments for improvement. This is because grades make them compare themselves, whilst written comments based on the task make them think about how to improve the work and receive the message that they can do better. In addition, marks only do not give any information about current achievement and necessary improvement. Comments only may support further learning.

Of course one cannot be sure that student peer feedback is always correct as teacher feedback. However, student peer feedback is readily and more immediately available than that of the teacher's. Research undertaken by Topping (2009) provided evidence that once learners become familiar with the process of peer-assessment (after time and adequate practice), reliability is likely to be quite high. Topping (2009) showed that a peer assessor, with less skill in assessment but more time in which to do it, can produce an assessment of equal validity and reliability to that of a teacher. Along the same line, Falchikov and Goldfinch (2000) carried out a meta-analysis of forty-eight quantitative peer assessment studies that compared peer and teacher marks. They concluded that students are able to make reliable judgements because peer-assessments were found to be similar to teacher assessments when students made judgements based on explicit criteria that they had well understood and had taken ownership of them. In Taiwan, Tseng and Tsai (2007) found out that peer assessment grades were highly correlated with teacher scores. Yet, the study by O'Donnell and Topping (1998) provided evidence that peer assessment can be as effective as teacher assessment and sometimes more effective.

Research on Teacher Assessment Practices

While the general research interest in classroom and formative assessment is high, research on teachers' competencies of assessment practices showed that in-service teachers' assessment knowledge and assessment strategies are limited. For example, Campbell and Collins (2007) and DeLuca (2012), based on documentation of teachers' classroom assessment practices, concluded that they are not skillful in assessment. The same researchers made the point that teacher education needs to give more attention to the development of classroom assessment strategies by teachers.

Brookhart and her colleagues (2016) strongly argued that teachers need help to improve the assessment criteria they use to grade, they need to be able to develop assessment criteria and then to effectively incorporate instructional skills with assessment skills. Brookhart (2011), not only argued in favor of the need for educators to give much attention to the development of assessment knowledge and skills for teachers, but she also proposed a list of assessment skills informed by current advances into research on classroom assessment.

Research on peer-assessment in Turkey

In Turkey, Acar-Erdol and Yıldızlı (2018) aimed to identify the classroom assessment practices implemented by instructors at the primary, secondary and high school level. By using Classroom Assessment Practices Survey, they found out that teachers had adopted an approach of assessment for learning and provided strong evidence of the advantages for students' learning. Secondly, Özdemir (2016) worked with prospective teachers to firstly give them the opportunity to practice peer-assessment and then, investigate the participants' views and opinions about peer-assessment. The participants talked extensively about the importance of providing constructive feedback and not grades. Again, from the teachers' perspective, Gelbal and Kelecioğlu (2007) provided evidence of the teachers' need to have practice and initial education in assessment since they identified problems in practice. In addition, Koç (2011) conducted a research study to document prospective teachers' opinions about peer-assessment in teaching practice. The participants talked about a wide range of benefits for both students and themselves as prospective teachers. Ozmen and Aydın (2015) conducted a research study within a teacher education program in Turkey with the aim to examine student teachers' beliefs about oral corrective feedback. Although the subject field was that of language learning and teaching (different from mathematics and physics), the findings are of significant importance for this study. Their evidence showed that although most participants held constructivist beliefs about teaching and learning, their strategies related to corrective oral feedback varied in terms of correcting errors. They reported on such terms as language proficiency, language components and task type.

Sasmaz Oren (2012) investigated the impact of gender and previous experiences on the approach of self- and peer-assessment in Turkey. Interestingly, they found out that female students received significantly higher mean scores than male students. Kayacan and Razı (2017) investigated the impact of self and anonymous peer feedback on four written assignments among high school students. Their findings strongly supported the idea that both self and peer feedback enhance the improvement of assignments in various ways.

In higher education, a study designed and conducted by Şahin (2008) looked at peer-assessment. The aim was to investigate peers when giving scores on project work. It was found that scores by peers were similar to those given by the instructor. In a following study, Şahin and colleagues (2016) looked at peer assessment of undergraduate students to report that those with high levels of achievement differed significantly from those made by students with medium or low level of achievement. More recently, Boztunç Öztürk and colleagues (2019) worked with 66 university students and focused on the analysis and comparison between analytic and general impression scoring in peer-assessment. One main finding was that students were distinguished from one another at a highly reliable rate using various scoring methods.

Method

Based on such research evidence, we, as teacher educators and university teachers, should think about how we develop and provide opportunities for prospective teachers to develop assessment practices in initial teacher education. Consequently, the focus of this research is on how prospective mathematics and physics teachers experienced peer-assessment during initial teacher education. More, specifically, the aim of the study is to explore the difficulties and challenges that prospective mathematics and physics teachers experienced while giving feedback and grades to their peers' coursework.

Research Design and Questions

A qualitative research study was designed to investigate the difficulties and challenges that prospective mathematics and physics teachers experienced while giving feedback and grades to their peers. The approach of the study is that of a case study (Stake, 1995; Yin, 2017) in an effort to collect in-depth and comprehensive information. Stake (1995) defined a case as having specific boundaries in terms of one phenomenon, time and place. For the specific research purposes, the boundaries are the teaching context of the course and the University context within which the research was conducted (as explained in the "Context of the Study" section) .

The study was designed to answer one main research question: "*How did prospective mathematics and physics teachers experience peer-assessment?*" with the following two sub-questions:

1. *What difficulties and challenges did prospective mathematics and physics teachers experience? (when giving feedback and grades to their peers)*

2. *How did such experiences change during one academic semester?*

It was not within the scope of the study to compare the instructor's feedback and grades with those of the participants. In addition, we are not interested in evaluating the assessment course, within which this research study was conducted. With regard its contribution, the present study aims to highlight the importance of including assessment education and development of assessment knowledge and skills in initial teacher education mainly in initial Mathematics and Physics teacher education.

Participants

Four cohorts of prospective mathematics and physics teachers in their third-year or fourth-year participated at a state university in Turkey. In total, ninety-four (94) undergraduate students were involved (during four consecutive academic semesters: Spring academic semester 2018 - Fall academic semester 2019). They were from the whole range of educational achievement, approximately of equal numbers of mathematics and physics prospective teachers. From the official records, it seems that they were 20-22 years old. In addition, 62% of them were women and 38% men. On average, prospective teachers from the four different academic semesters are samples drawn from the same population. For all of them, Turkish was their native language with English as their first foreign language and official language of teaching at the university. Purposive sampling (Creswell, 2013) was applied, whereby the researcher intentionally selects the participants to investigate the research problem. The names of the prospective teachers have been removed. Instead, we have used numbers for each participant (1-94), as the following Table shows.

Table 1 Cohorts and Participants

	Mathematics	**Physics**	**Total**	**Numbers**
1st Cohort Spring Academic Semester 2018	10	12	22	1-22
2nd Cohort Fall Academic Semester 2018	14	12	26	23-48
3rd Cohort Spring Academic Semester 2019	13	11	24	49-72
4th Cohort Fall Academic Semester 2019	12	10	22	73-94
Total	49	45	94	

The Context of the Study - The Classroom Assessment Course

The study was carried within the context of a classroom assessment course for 13 full weeks, for each of the four cohorts. The course was developed and taught by the main research investigator, who was the instructor of the four cohorts, too. It was the first compulsory course about classroom assessment for prospective teachers. The participants were introduced to formative and summative assessment, the development of rubrics and grading, feedback and peer-assessment. Further details about the syllabus of the course are given in Appendix A.

The participants were giving oral and written feedback and grades on coursework during one academic semester. Once a week, after the submission of coursework, they would take one peer's piece of coursework as homework (out of class) to give written feedback in the form of comments and grades. Although comments are of major importance, the participants had practice in grading, because grading is a key professional skill for teachers. In addition, one of the main aims of the assessment course was them to develop a sound understanding of grades and grading in support of learning. The instructor would need to look at how well each of them did so that matching of students with coursework was deliberate and selective, according to the achieved coursework. They would bring it back to the class and discuss in pairs for around ten minutes. The participants would talk to each other in groups of two (two different groups for each participant; as assessor and as assessee). For example, if there were 24 prospective teachers in the class, 24 conversations would take place in pairs. Each participant (the assessor, in each pair) gives feedback and they discuss the piece of coursework. Then, they change pairs so that the assessed peer becomes the assessor, who initiates a discussion and gives feedback in a different pair for ten more minutes. They explain the written feedback and they give oral feedback, too. The intention is to identify strengths and strong points, as well as to offer hints for improvement when weak points exist. After peer-assessment, time was given for revision so that each student improved the paper and submitted it again. However, this paper reports only on the peer-assessment process. A different paper has focused on self-assessment experiences by the same prospective teachers. All discussions were recorded with microphones put in front of the participants so that neither the instructor nor the two research assistants interrupted or distorted the natural conversation of students.

Methods of Data Collection

The study employed rigorous data collection involving:

- collection of weekly reflective journals written by the participants,
- semi-structured interviews and,
- collection of written feedback by peers to each other.

Reflective journals (Boud, 2001) were submitted once every week. The participants were asked to write reflective papers in order to talk about their experiences of peer-assessment. They were not asked any particular questions because we would not like them to get biased by our questions. Instead, we wanted them to select the issues they were concerned with (grades, feedback and whether the process helped them learn better). For example, one expression was: "*In your journal, write down how you have experienced peer-assessment this week. What are your experiences when giving and receiving feedback*?" Boud (2001) argued in favour of journals promoting reflection on practice. For Boud, journal writing should capture and enhance reflection on professional practice and learning.

The role of journal writing is to give an account of what happened and to retrieve as fully as possible the rich texture of events as they unfolded (Boud, 2001, p. 14).

Reflective journals were kept private. This gives freedom and confidentiality so that each participant expresses herself and provides a sincere flow of thoughts and feelings. From the beginning of the course, it was clarified that reflective journals were not going to be graded.

Semi-structured individual interviews (Kvale & Brinkmann, 2009) were carried out to shed light on the participants' experiences and difficulties. An interview protocol was developed for the purposes of this study and it was piloted. While the open questions in the list formed the main impetus within the interviews, we were also open to probing ideas that they were raised within the discussion, particularly if the questions seemed important to the participants. Appendix B gives the interview questions. Interviews were carried out individually. Each one lasted approximately for 10-15 minutes. All interviews were audio-taped with the participants' written consent and transcribed. Lastly, written feedback given by the participants was collected for analysis.

Through the three different research methods, experiences are rarely missed but rather collected at different times so that the whole picture can be viewed rather than considering only snapshot examples. Thus, what they have written in reflective journals can be elicited later in the interviews. In addition, the analysis of the feedback (oral and written) they exchange can provide the basis for interviews so that different perspectives can add to the richness of the data. Finally, all participants were informed about the research purposes of the study, gave their written consent and ethics guidelines were kept.

Data Analysis

The analysis of data was an inductive process of narrowing data into a range of emerging themes and trends (Creswell, 2013) through multiple read of the data. Open coding techniques and triangulation were applied. To make sense out of text data, the text was divided into segments, different segments were identified and the segments were labelled with themes that accurately described the meaning of the text segment. Similar themes were aggregated together to form a major idea until no new themes would emerge (Silverman, 2001). We stopped developing themes when we identified the major themes and no new information could be added to our list of themes or to the detail for existing themes. Thus, themes describing the participants' experiences emerged from the data.

In addition, greater confidence in our findings was developed because the analysis also focused on an extensive use of "triangulation" (Creswell & Miller, 2000) to provide several viewpoints from different participants and sources of data as sufficient evidence for each emerging theme. Triangulation was essential to consistently validate and verify data and research findings. With regard to the second research question, that is to document likely changes of each participant during the study, interview answers and reflective journals were compared either from different participants, or from the same participant at different times in the study (Creswell, 2013).

The interview questions were validated in two ways. Firstly, they were piloted with some undergraduate students before the actual study started (with a different class than the four cohorts which participated in the main study). Secondly, an expert at assessment checked the questions. The principal investigator and the two research assistants analyzed the data. The level of interrater agreement between the main investigator and two researchers was 82% during the coding process, with discussions held in order to reach consensus on final decision. All participants agreed that the quotations as selected and presented in this paper give a fair representation of their experiences and views.

Findings and Discussion

The main findings are presented under the headings of five themes:

Table 2 Main Findings – Main Themes

Theme 1: Grades are more important than feedback because they show what has been achieved. Yet, deciding on grades is more difficult than giving feedback.
Theme 2: Written feedback (in the form of comments) is more important for learning than grades.
Theme 3: Participants' experiences of giving oral and written feedback.
Theme 4: Restricted quality of written feedback.
Theme 5: Feedback related to the assessment criteria. The focus on feedback to show ways for improvement of coursework.

Quotations from the interview answers, reflective journals and the written feedback are provided to illustrate the findings. After each quotation, the reader will find a parenthesis in which there will be one number for each participant (because names have been removed for anonymity purposes), the number of the cohort, the number of the week (1-13), the subject of each participant (Mathematics or Physics) and finally, the data source (interviews or reflective journals). Lastly, any grammar or spelling mistakes in students' quotes are being preserved.

Theme 1: Grades are more important than feedback because they show what has been achieved. Yet, deciding on grades is more difficult than giving feedback.

In the very first and second weeks, the majority of prospective teachers reported that peer-assessment was a unique experience, which they enjoyed. The following statement is representative of such experience:

I think reading and assessing the coursework is exciting and interesting work (8, cohort 1, week 1, Mathematics, reflective journal).

They also explained that peer-assessment, by giving comments and grades, made them confident as they started feeling to be trusted by the teacher and peers. The feeling of being trusted was important and it was further developed during the academic semester. For example:

I have liked being important and being trusted because for all my education years, for much important modules, I've never assessed someone else's coursework. Because in my mind I believed that grading was everything and students or peers could not be trusted, only a teacher can do this (64, cohort 3, week 2, Physics, interview).

In addition, the majority of participants wrote in the reflective journals and explained in the interviews that the provision of feedback is not sufficient. This is because it is mainly grades which inform students about where they are and how they have done. Grades are needed for motivation, too. One representative quote is the following:

Feedback only is not sufficient. Teachers need to give grades because students need to know how they have done. It is grades which give such information (32, cohort 2, week 2, Mathematics, interview).

This is an interesting aspect, which further explains their worry about giving a grade to a piece of coursework. In fact, they considered grading to be more difficult than offering comments. One main reason they gave was that grades are more important for students' success.

At the same time, giving comments was easier because comments do not count for the final grade or for the GPA. Typical statements of such thinking were as follows:

Writing comments was simpler than giving grades in peer-assessment (77, cohort 43, week 1, Mathematics, interview).

The crucial role of grades for one's success made them be careful and reluctant to give grades. One mathematics and one physics teacher, respectively, explained in the interviews:

I feel uncomfortable, and this is the hardest of the two when I should give grades. Giving written feedback like comments is easier for me (81, cohort 4, week 2, Mathematics, interview).

I gave comments to my classmate. However, it was hard to decide what grade to give to my peer. I have changed my decision for a few times (45, cohort 2, week 2, Physics, interview).

Thus, although they enjoyed looking at one peer's piece of work, they wanted to avoid grading. The explanations the participants gave in the interviews help us create a better picture of the limited understanding of feedback and grades and, then, of peer-assessment they have developed in the first weeks of the course. One physics participant stated: *Because you are not allowed to make a mistake. Grades are important for your GPA* (89, cohort 4, week 2, Physics, interview).

Theme 2: Written feedback (in the form of comments) is more important for learning than grades.

During the course, they have practiced peer-assessment whereby they had the chance to offer and receive written comments as feedback and oral feedback, too. Such practice resulted in significant changes in participants' thinking for the majority of them. They got more actively engaged in the process, which made them consider the importance of feedback in the form of written comments for improvement of the coursework. The ideas with which they joined the study changed towards a better understanding of the role of peer-assessment. With regard feedback, many participants started understanding that the issue is not to give grades only but mainly advice about how to improve a piece of coursework.

Giving comments is a much higher skill than giving grades (91, cohort 4, week 3, Physics, reflective journal). And,

I think the most difficult part of the assessment is giving feedback to my peer (54, cohort 3, week 3, Physics, interview).

It has been clear that they changed their views related to grading, as they practiced giving feedback to peer's coursework.

This assessment activity gave us a chance to improve our work. In the case that we only obtained our grades, there would be no improvement. However, feedback from our peers helped us improve our work (69, cohort 3, week 4, Physics, interview).

All these show that there is a clear transition from an initial idea that grades and grading are the most crucial to the idea that written comments are also needed because they provide information about achievement and quality of work. Many times, they wrote in the reflective journals:

Before I got this course, grading was more important, but now feedback (88, cohort 4, week 4, Physics, reflective journal).

Theme 3: Participants' experiences of giving oral and written feedback.

When the participants returned the coursework with written feedback and grades, they had discussions in pairs to discuss written feedback. They reported a range of experiences and benefits from such discussions whereby they shared feedback. The following statement is clear:

We have talked in two different pairs. My classmate asked me a few questions. In my turn, I have asked the points that I did not understand to the person who assessed my piece of coursework (8, cohort 1, week 9, Mathematics, interview).

Such discussions helped them clarify any likely points that were not well understood and make corrections in a supportive environment. Many times they reflected on such discussions they had with peers in pairs. The following statement is representative of the clarification purpose of discussion in pairs:

In this peer-assessment activity, discussions were helpful because we tried to convince each other about our feedback and work. Or, if there are some points that we did not understand well, we had a chance to discuss the point, the topics, as well (51, cohort 3, week 9, Mathematics, interview).

Peer-assessment have led to discussion and maybe disputes between prospective teachers. In such discussions, they learn better and more because they discuss in equal terms. This is because there is no peer who holds any authority (as it is the case with the "traditional" classes in which only the teacher has authority to make decisions). In addition, as classmates, they communicate in a language that they naturally use. A trusting and a more relaxing relationship has been developed, as they have explained with the following quote:

Discussion with our papers is also helpful because we use the same basic language (78, cohort 4, week 9, Mathematics, interview).

Talking with classmates enables us to re-think about our written work and improve it. We listened to classmates' opinions and why s/he thinks like that. If we get only grades probably we may not be able to improve our work (10, cohort 1, week 5, Mathematics, reflective journal).

Thus, it makes sense that for many of them, such discussions were one of the best parts of the whole process, as the following quotes state:

The best part of the peer-assessment is giving and taking oral feedback from my classmates, communicating with them. I learned from him and also I saw my weaknesses (29, cohort 2, week 5, Mathematics, reflective journal) and,

Talking about peer-assessment with friends is very important because we have the chance to ask: "Why did you give me this grade?" and "How do I improve my work?" and discuss these issues (82, cohort 4, week 5, Mathematics, reflective journal).

All these quotations make clear that participants' discussion in pairs helped them understand the feedback and how they need to revise their coursework. During the semester with practice, they would no longer worry about the peers who were classmates and friends. They enjoyed discussing with them, giving feedback and receiving guidance.

Theme 4: Restricted quality of written feedback.

Under this theme, we have included feedback which was of limited information for the participants. This means written feedback which was restricted to general comments, identification of mistakes and confirmation of right answers. The analysis of written feedback showed that the majority of participants, when started to give feedback, they wanted to confirm correct answers, identify mistakes and give general comments like the following: "*This is a good job*", or confirm right answers like "*I like your answer*" or send a message like "*This answer is wrong"*. In particular, they gave general comments*: Very good. Good work. Nice job. Excellent work. Good effort. Vague. This is a great answer!*

They gave vague recommendations: *More is needed here. Improve this answer. You are not clear here. I do not see what your point is here. And sometimes they gave general advice: Study* *more! This needs improvement. Be more careful. Do your best. Study more. You forgot to give an example. You need more details in your answers.*

Theme 5: Feedback related to the assessment criteria. The focus on feedback to show ways for improvement of coursework.

The participants carried on improving their understanding of feedback. In particular, they realized that for feedback to fulfill its specific role to support further learning, it should be related to the assessment criteria. It should, then, identify the strengths and weaknesses (in relation to assessment criteria) and propose ways for improvement. Under this theme we have included types of feedback which are closely related to the assessment criteria and suggestions are explicit. In addition, sometimes such feedback may be put in the form of a question to make the recipients think about how to respond to it. All the following quotations show how feedback can be specific to the subject (Mathematics and Physics) and, to the task; what this study has called subject-specific feedback. The following comments illustrate the high quality subject-specific feedback that prospective teachers gave to their peers (mostly from the mid-term until the end of the academic semester).

Please refer to your graph when you analyze your data.

What about talking about how you will improve your experiment, if you repeat it.

How do you explain your experimental evidence by using your theory?

You forgot to talk about limitations in your experiment.

After you plotted the data points, how have you drawn the curve?

What do you think about the forces acting upon the mass? Can you sketch them?

I like the way you organized your ideas and examples in your answer.

This is a complete answer because you have included all the possibilities of different line graphs.

Think about what the learning targets are for today's lesson.

When reflecting on the process of feedback they wrote:

Peer assessment is a good way to let your classmates know about their weaknesses and strengths because we are in the same class (70, cohort 3, week 5, Physics, reflective journal).

I was very pleased while I am/was examining my peer's paper. Actually, peer-assessment is very beneficial because when we write our papers, we cannot see where we have any *drawbacks but when we assess our friend's work, we can easily see the strengths and weaknesses of her paper* (54, cohort 3, week 5, Mathematics, reflective journal).

The above quotations provide much information about the process and its benefits to both parts (assessors and assesses). In fact, it is easier for anyone to identify mistakes and weak points in the peer's work than in his own and, then, come back to his/her own paper and make the relevant corrections. At the same time, the assessor, when looks at a better piece of work, can take ideas from strong points and improve his/ her own work. Also, we need to underline that the participants (by time and practice) have supported each other, they have collaborated in order to learn better and improve their work. There are no longer worries about how not to offend classmates and so on. They have come through a range of learning experiences in which they developed confidence that together they can learn better. They take time to use feedback and improve their work as the next quotations show:

With my peer's comments, I will improve my answer and learn things that I do not know (34, cohort 2, Mathematics, interview).

Peer-assessment has given new ideas to me. I looked at my own answers from my classmate's different perspective (39, cohort 2, Mathematics, interviews).

Thus, peer-assessment is a two-way process. Each participant has helped his/her peer with the written feedback they offer and at the same time, they get ideas about how to improve their own coursework. Finally, they underlined the learning aspect of the process and their attention to the assessment criteria:

What else is needed here so that you explain your ideas better? (Hint: think about the assessment criteria) (written feedback)

Since we learned about the assessment criteria in the lesson, applying them when looking at the work of our classmates was a great learning experience (11, cohort 1, week 12, Physics, reflective journal).

"*I learn better by giving feedback*" (47, cohort 2, week 12, Physics, reflective journal).

This is very important. This is true because when they give feedback, it is easier for them to identify mistakes in the peer's work more easily than in their own. Many prospective teachers wrote:

I found my friend's coursework better than mine. I learned from his paper (52, cohort 3, week 12, Physics, reflective journal).

The real learning happened when I evaluate the strengths and weaknesses of another person [...] *When you realize your strengths and weaknesses, you can improve your work in a wonderful manner* (77, cohort 4, week 12, Mathematics, reflective journal).

The meaning of "wonderful" is to do with immediate and explicit ideas for improvement and the fact that they have collaborated with their classmates with whom they speak the same language. They have talked explicitly about the process of understanding the assessment criteria better and applying them. The participants discuss the assessment criteria and then they assess peer's coursework. The process of giving subject-specific feedback and discussing in pairs, resulted in participants' understanding and learning the assessment criteria and what is required for a high quality work.

For most of them the process was interesting since they would need to think about the expected quality of answers and focus on the related assessment criteria.

... applying the assessment criteria to assess our classmates' work was a great learning method…. (89, cohort 4, week 12, Physics, interview) and,

Since we discussed the assessment criteria of our written work, applying them to assess our friends' work was a very helpful experience. Now I much better understand how some work is assessed. The process was interesting (27, cohort 2, week 12, Mathematics, reflective journal)

Having gone through the practice of the peer-assessment process, they concluded*:*

Peer-assessment is an intense and time-demanding process. It requires, firstly, good understanding of the topic to enable you to assess effectively. Secondly, it requires time to go through and assess the work and time to give constructive written feedback (7, cohort 1, week 13, Mathematics, interview).

By participating in peer-assessment, I learned many things: how to assess, how to give feedback and how to discuss our feedback with our peers (52, cohort 3, week 13, Mathematics, reflective journal).

Conclusions and Suggestions

The practice of peer-assessment has opened the way to a wide range of experiences for prospective mathematics and physics teachers. As they wrote in the reflective journals and explained in the interviews, they did like the process of giving comments and grades. It was the first time they were given such an opportunity; this was, so far, solely the instructor's duty. They experienced the feeling of being powerful to give grades and being trusted by the teacher and classmates. First of all, the findings, as already presented in the previous section, agree with those by similar studies on assessment in Turkey (i.e. Gelbal & Kelecioğlu, 2007; Koç, 2011; Özdemir, 2016).

One main finding of the present study is that in the first weeks of the course, prospective teachers demonstrated a limited understanding of peer-assessment due to a limited understanding of comments and grades. The written comments they started giving were restricted to general ones: "*This is a good job*", or to confirm right answers like "*I like your answer"* or to identify wrong answers "*This answer is wrong*". They also supported the view that grades are more important for students, because grades do matter for success. In addition, the participants supported the view that grades offer better information about student achievement than feedback does. Thus, the majority of them appeared to give simple comments and also not to feel comfortable to assign grades. For them, writing simple comments does not require much time and effort; it is easy and simple. In addition, according to them, giving grades requires more thinking and effort. Taking it as a whole, their understanding of peer-assessment seemed to be limited as it is clear that they did not realize the full potential of feedback; for how to improve a particular piece of coursework and for how to support current and future learning. This very first finding of the present study seems to be in contradiction with what research has shown. In the literature review section, it was made clear that Butler (1988), among more researchers, provided strong evidence that marks and feedback (when provided together) do not help students improve their work. This is because they get distracted by marks and thus, they do not pay any attention to the comments offered to them for improvement. Or, grades only, do not contribute to learning because they do not give them any information about current achievement or further learning. However, studies like the one carried out by Smith and Gorard (2005) reported on an intervention in which they did not give any grades to students but only comments to help them understand what they needed to improve. Although, according to the intervention, students should have focused their attention only on comments, the participants claimed their marks because it was marks that they could understand better.

On the other hand, it seems that there is one paradox here. The participants' idea that giving feedback is more straightforward than giving marks reflects a naive idea, which ignores the crucial role of feedback in the learning process. "Writing comments was simpler than giving grades in peer-assessment". While prospective teachers were not familiar with feedback in the form of comments, but with grades only, they stated that giving feedback was easier than giving grades. They seemed to hold a naive idea of comments, not related to the role of feedback in support of learning. Interviews shed more light to such experiences and perceptions so that we better understand their reasoning. Such an experience may be interpreted by considering that until that time, their experiences were only related to summative assessment and grades. Indeed, they joined the course with "initial" ideas that marks are important for one's success and for the total GPA. The dominance of summative assessment resulted in a limited understanding of peer-assessment because their experiences originated from grade-dominated teaching. Secondly, peer-assessment was carried out in the class environment with 22-26 prospective teachers with classmates as peers. This resulted in them not feeling comfortable to do peer-assessment by giving grades to classmates with whom they were friends.

Through the development of the course, many of the participants started experiencing that grades by themselves cannot give any information about further learning. When they realized what the meaning and the role of feedback are for the learning process, they emphasized that giving feedback is more important. Many of the participants were concerned about good quality feedback and not concerned about grades. They started recognizing the value of the peer-assessment approach. They recognized that feedback should not focus on the person and this delicate notion helped them proceed when the peer was their classmate. This makes sense because the classes, in which the study took place, were not large. The development of an appropriate understanding of the role of feedback would take a whole academic semester. In addition, the present study provided strong evidence that there was a development in the type of feedback they gave to each other. From the third week onwards, their experiences shifted towards better peer-assessment practice. In parallel with the progress of the assessment course, they experienced that they can have a role in the assessment process; this idea could not be shared by lecturing only. They started developing a better understanding of the role of feedback in the learning process. They started adopting research-based assessment strategies; mainly that feedback is more important for one's learning than grading.

It took them enough time to start thinking in terms of the assessment criteria and the required quality so that written feedback to be related to the assessment criteria. It required also much effort for them to be convinced that firstly, they need to think about the provision of good feedback and then, about grades in support of learning (Brookhart, 2017). Their understanding of feedback improved when they started experiencing that feedback (in relation to assessment criteria and the required quality) is more crucial for one's learning. Thus, feedback should not be restricted to show right or wrong answers but to identify strong points, weaknesses and missing points. Feedback should then show ways for improvement. In other words, feedback should enhance learning by highlighting the strengths and weaknesses and also by giving suggestions about what needs to be done so that coursework will get improved.

Discussions with peers in pairs helped them advance their understanding of assessment criteria and subsequently, of feedback. Discussions seemed to be an intermediate “step” between simple types of feedback (Theme 4) and feedback related to the assessment criteria (Theme 5). The participants explained how the discussions in pairs helped them improve coursework and learn better. This was because the discussions had a real purpose: for them to have the opportunity to communicate written feedback, to clarify likely difficult points and suggest ways for improvement. Prospective teachers considered it as the most valuable part of the peer-assessment process, because they felt comfortable to talk to their peers about their coursework. In fact, they were much more comfortable than talking to the teacher. They would ask questions to their peer more easily than to the teacher. They take responsibility for their own learning on an equal basis with their peer. The participants in this study enjoyed discussing on the basis of their work as they would talk and use the same natural language. They see that they can take some responsibilities and make decisions about their peer’s learning, which, so far, was the responsibility of the teacher only. They have worked within the zone of proximal development to “scaffold” peer’s learning (Vygotsky, 1986). Through peer interaction, successful peer feedback is being communicated to support the learning process. In discussions, prospective teachers attend to each other as resources for learning (Wiliam & Leahy, 2015). The process becomes transparent because they are able to understand where they are in their learning, where they need to go and how best to get there (Sadler, 1989). Learning how to give and accept comments, how to make suggestions, how to justify one’s position and reject suggestions are all important skills. A full understanding of feedback is achieved when the aim for each of them is to advance the quality of peer’s coursework. They take time to use feedback and improve their work. Such an experience is important because the theory on classroom assessment highlighted the need for the feedback to be understood by students and then, to be used for improvement of coursework and for a second submission (Black & Wiliam, 1998). Their engagement in applying the assessment criteria resulted in them giving subject-specific feedback, which by itself, is important knowledge for prospective mathematics and physics teachers.

Notably, there is one more important issue related to the practice of giving written feedback. Feedback seems to be immediate; in a short time after coursework submission. The participants would not need to wait for a certain time as it is the case when feedback is provided by the teacher. Teacher feedback is mostly not offered in time for students to benefit immediately. Right after discussions, they would get time to work on it and improve their work. The whole process, although it may seem to be time-consuming, it is not. Much time was saved. But in peer assessment feedback is given immediately by peers so that they are able to act upon it and improve their own work.

The progress that prospective teachers made in giving feedback (from general comments to subject-specific comments) made them more critically aware of the process of peer-assessment. The participants experienced that comments guiding the learner toward the next step are more valuable, as they promote learning. Consequently, the idea and confidence that they can improve attainment is apparent. All of them participated to do their best. They were actively involved in applying the assessment criteria and demonstrating the expected quality they would need to demonstrate in the coursework. Thus, they explained that this sort of peer-assessment helps them shift the emphasis towards improvement of coursework. They developed the idea and confidence in themselves that they can improve attainment, and then how to help their peer.

Giving feedback and grades requires, first of all, good subject knowledge (they may need to go back to their notes) and development of assessment knowledge. Consequently, they gained confidence in providing feedback in their own subjects, Mathematics and Physics. In fact, they gave good quality feedback to their peers; both written and oral (when discussing in pairs). This experience has been rewarding: to write comments on strengths and weaknesses and give ideas about how to improve their peers' and their own coursework. The whole process is double-sided because both parts have benefited: the assessor and the assessed peer. The emphasis was not only on the final product (the final version of coursework) but primarily on the process. They also learned that good quality feedback is specific to the subject (Mathematics and Physics), the specific tasks and coursework.

The present study provided a range of challenging experiences with regard classroom assessment in initial teacher education. It also provided strong evidence that the course improved their assessment knowledge and assessment skills to some extent. On the basis of such research findings, we want to argue that the development of assessment knowledge and research-based assessment skills does take much time. In fact, more practice is needed as there was not enough time for them to develop skills towards the direction of grading. The duration of 13 weeks was not enough. A second assessment course is needed. In fact, more time is needed for the prospective teachers of the present study so that they develop an adequate understanding of grading in support of learning.

This study highlights the importance of including assessment education and development of assessment knowledge and skills in initial teacher education. Shepard and her colleagues (2005) strongly argued that prospective teachers should be taught about assessment which supports learning by practicing certain research-based assessment methods and giving feedback to their own work. The study, by its design, did not attempt to generalize findings to all prospective teachers. Such an aim would be beyond the context of the present study. We should also consider the constraints of one academic semester. We would like to carry on the analysis of the data in order to compare the participants' comments and grades with those of the instructor's.

Understanding of prospective teachers' experiences, difficulties and challenges during their practice is necessary for the design of effective initial teacher education programs. Prospective teachers experience difficulties and have needs which must be considered when planning and implementing initial teacher education. Prospective teachers need help to move from the focus on grades only, towards the learning of assessment criteria and finally, to focus on written feedback - comments related to assessment criteria. Prospective mathematics and physics teachers need support to apply assessment criteria to the specific subjects of Physics and Mathematics. The role of the Departments of Mathematics and Physics Teaching in developing assessment practices in prospective teachers is very important.

Further research is needed to follow the same participants when they obtain teaching posts in schools to identify the range of needs they experience in the first years of teaching in schools. Another suggestion would be to continue working with the same participants to further develop classroom assessment skills. Last, but not least, we would like to invite Mathematics and Physics educators to investigate similar classroom assessment issues with prospective mathematics and physics teachers at different universities.

Acknowledgements

Special thanks to all prospective teachers who participated in the study and gave much effort and time from their busy schedule. To the two research assistants who helped with data analysis.

Appendix A

Main Topics of the Assessment Course

Classroom Assessment

Development of learning targets

Development of assessment goals and criteria

Development of assessment tasks

Validity and reliability of assessment

Alignment among curriculum, instruction and assessment

What makes feedback effective?

Feedback and Grades: What does research say?

Assessment for Learning

Grading- Using learning targets to guide summative assessment and grading

Performance Assessment

Appendix B

Questions in semi-structured interviews

What is your experience with peer-assessment?

Did you like it?

Why? Yes? Or why not?

What is it that you liked?

What is it that you did not like?

What did you learn?

Did you help your peer? How?

Did your peer(s) help you? How?

What was difficult for you?

Was giving feedback difficult? Why?

Was giving comments difficult? Why?

How did it help you? Was giving grades difficult?

What did you learn? Please give me an example.

Are you going to use peer-assessment when you are teacher in a school?

What sort of classroom assessment have you found important?

How confident are you in developing assessment criteria?

How did feedback help you find out how to do better? How did peer-assessment help you?

Matematik ve Fizik Öğretmen Adaylarının Akran Değerlendirmesini Uygulama Deneyimleri

Özet: Sunulan araştırmanın amacı, matematik ve fizik öğretmen adaylarının akran değerlendirmesini nasıl deneyimlediklerini araştırmaktır. Çalışma, hizmet öncesi öğretmen eğitimi için tasarlanmış bir sınıf içi ölçme-değerlendirme dersi kapsamında gerçekleştirildi. Çalışmaya dört gruptan doksan dört (94) öğretmen adayı katıldı. Katılımcılar akranlarının ödevlerine geri bildirim verdiler, yansıtıcı günlükler yazdılar ve yarı yapılandırılmış görüşmelere katıldılar. Katılımcılar, not vermenin geribildirim vermekten daha zor olduğunu, çünkü notların öğrencilerin başarısı için "sayıldığı" için önemli olduğunu bildirdi. İkinci olarak, ders boyunca katılımcıların düşünceleri, yorum vermenin not vermekten daha önemli olduğu fikrine "kaydı" çünkü geribildirim onların değerlendirme kriterlerini anlamalarına ve ardından dersleri iyileştirmelerine yardımcı oldu. Bulgulara dayalı olarak, makalenin ana argümanı, öğretmenlerde sınıf içi değerlendirme becerilerinin ve değerlendirme bilgilerinin geliştirilmesinde hizmet öncesi öğretmen eğitiminin rolü hakkındadır.

Anahtar kelimeler: değerlendirme kriterleri, geri bildirim, not verme, akran değerlendirmesi, hizmet öncesi matematik ve fizik öğretmen eğitimi.

References

Acar-Erdol, T., & Yıldızlı, H. (2018). Classroom Assessment Practices of Teachers in Turkey. *International Journal of Instruction, 11*(3), 587-602. https://doi.org/10.12973/iji.2018.11340a

Black, P. (2003). *Testing: Friend or Foe? Theory and Practice of Assessment and Testing*. Routledge Falmer Press.

Black, P. & Harrison, C. (2004). *Science inside the black box: Assessment for learning in the science classroom*. King's College and GLAssessment.

Black, P. & Wiliam, D. (1998). Inside the black box: Raising standards through classroom assessment. *Phi Delta Kappan, 80* (2), 139-144.

Boud, D. (2000). Sustainable assessment: Rethinking assessment for the learning society. *Studies in Continuing Education, 22,* 151-167. https://doi.org/10.1080/713695728.

Boud, D. (2001). Using journal writing to enhance reflective practice. *New Directions for Adult and Continuing Education, 90,* 9-17.

Boztunç Öztürk, N., Şahin, M.G., & Ilhan, M., (2019). An analysis of scoring via analytic rubric and general impression in peer assessment. *Turkish Journal of Education, 8*(4), 258−275. DOI: 10.19128/turje.609073

Brookhart, S. M. (2011). Educational Assessment Knowledge and Skills for Teachers. *Educational Measurement: Issues and Practice, 30*(1), 3-12. https://doi.org/10.1111/j.1745-3992.2010.00195.x

Brookhart, S. M. (2017). *How to Use Grading to Improve Learning.* ASCD Publications.

Brookhart, S. M & McMillan, J. H. (2020). Introduction. In S. M. Brookhart and J. H. McMillan (eds.) *Classroom Assessment and Educational Measurement* (pp. 1-7). Taylor and Francis Publications.

Brookhart, S. M., Guskey, T. R., Bowers, A. J., McMillan, J. H., Smith, J. K., Smith, L. F., Stevens, M. T. & Welsh, M. E. (2016). A Century of Grading Research: Meaning and value in the most common educational measure. *Review of Educational Research, 86*(4), 803-848. https://doi.org/10.3102/0034654316672069.

Butler, R. (1988). Enhancing and undermining intrinsic motivation; the effects of task-involving and ego-involving evaluation on interest and performance. *British Journal of Educational Psychology, 58* (1), 1-14. https://doi.org/10.1111/j.2044-8279.1988.tb00874.x

Campbell, C. & Collins, V. L. (2007). Identifying essential topics in general and special education introductory assessment textbooks. *Educational Measurement: Issues and Practice, 26*(1), 9-18. https://doi.org/10.1111/j.1745-3992.2007.00084.x

Creswell, J. W. C. (2013). *Qualitative inquiry and research design. Choosing among five approaches* (3rd ed.). Sage.

Creswell, J. S. & Miller, D. L. (2000). Determining validity in qualitative inquiry. Theory into Practice, 39 (3), 124-130. https://doi.org/10.1207/s15430421tip3903_2

DeLuca, C. (2012). Preparing Teachers for the Age of Accountability: Toward a Framework for Assessment Education. *Action in Teacher Education, 34*(5-6), 576-591, DOI: 10.1080/01626620.2012.730347

Falchikov, N. & Goldfinch, J. (2000). Student peer assessment in higher education: a meta-analysis comparing peer and teacher marks. *Review of Educational Research, 70*(3), 287-322.

Gelbal, S., & Kelecioğlu, H. (2007). Öğretmenlerin ölçme ve değerlendirme yöntemleri hakkındaki yeterlik algıları ve karşılaştıkları sorunlar. *Hacettepe Üniversitesi Eğitim Fakültesi Dergisi, 33,* 135-145.

Hattie, J. (2012). *Visible learning. Maximizing Impact on Learning*. Routledge. Taylor and Francis Publications Group.

Hodgen, J. & Wiliam, D. (2006). *Mathematics inside the Black Box: Assessment for Learning in the Mathematics Classroom*. GLassessment.

Kayacan, A., & Razı, S. (2017). Digital self-review and anonymous peer feedback in Turkish high school EFL writing. *Journal of Language and Linguistic Studies, 13*(2), 561-577.

Ketonen, L., Hähkiöniemi, M., Nieminen, P. & Viiri, J. (2020). Pathways through Peer Assessment: Implementing Peer Assessment in a Lower Secondary Physics Classroom. *International Journal of Science and Mathematics Education, 18,* 1465–1484 https://doi.org/10.1007/s10763-019-10030-3.

Koç, C. (2011). The Views of Prospective Class Teachers about Peer Assessment in Teaching Practice. *Educational Sciences: Theory and Practice, 11* (4), 1979-1989. ERIC Number: EJ962684

Kvale, S. & Brinkmann, S. (2009). *InterViews : Learning the craft of qualitative research interviewing*. Sage Publications.

O'Donnel, A. M. & Topping, K. J. (1998). Peers assessing peers. Possibilities and problems. In K. J. Topping and S. Ebly (eds.). *Peer-assisted learning* (pp. 255-278). Lawrence Erlbaum Associates.

Ozmen, K. S., & Aydın, H. U. (2015). Examining student teachers' beliefs about oral corrective feedback: Insights from a teacher education program in Turkey. *Australian Journal of Teacher Education (Online), 40*(12), 141–164. doi/10.3316/informit.681001963960425

Özdemir, S. (2016). The opinions of prospective teachers on peer assessment. *Educational Research and Reviews, 11*(20), pp. 1859-1870. DOI: 10.5897/ERR2016.2997

Sadler, D. R. (1989). Formative assessment and the design of instructional systems. *Instructional Science, 18,* 119-144. https://doi.org/10.1007/BF00117714.

Sadler, R. (1998). Formative Assessment: revisiting the territory. *Assessment in Education, 5*(1), 77-84. https://doi.org/10.1080/0969595980050104.

Sasmaz Oren, F. (2012). The effects of gender and previous experience on the approach of self and peer assessment: a case from Turkey. *Innovations in Education and Teaching International, 49* (2), 123-133, DOI: 10.1080/14703297.2012.677598.

Shepard, L. A. (2000). The role of assessment in a learning culture. *Educational Researcher, 29*(7), 4-14. https://doi.org/10.3102/0013189X029007004.

Shepard, L., Hammerness, K., Darling-Hammond, L., Rust, F. with Snowden, J. B., Gordon, E., Gutierrez, C. & Pacheco, A. (2005). Assessment. In L. Darling-Hammond and J. Bransford, J. (eds.) *Preparing Teachers for a Changing World. What teachers should learn and be able to do* (pp. 275-326). San Francisco, CA: Jossey-Bass Wiley National Academy of Education.

Shute, V. J. (2008). Focus on formative feedback. *Review of Educational Research, 78* (1), 153-189. https://doi.org/10.3102/0034654307313795.

Silverman, D. (2001). *Interpreting qualitative data* (2nd edition). Sage.

Sluijsmans, D. M. A. & Prins, F. (2006). A conceptual framework for integrating peer assessment in teacher education. *Studies in Educational Evaluation, 32* (2006), 6-22. https://doi.org/10.1016/j.stueduc.2006.01.005

Sluijsmans, D. M. A., Brand-Gruwel, S. & van Merriënboer, J. J. G. (2002) Peer Assessment Training in Teacher Education: Effects on performance and perceptions. *Assessment and Evaluation in Higher Education, 27*(5), 443-454. DOI: 10.1080/0260293022000009311.

Smith, E. & Gorard, S. (2005). 'They don't give us our marks': The role of formative feedback in student progress. *Assessment in Education: Principles, Policy and Practice, 12*(1), 21-38, DOI: 10.1080/0969594042000333896

Stake, R. E. (1995). *The art of case study research*. Sage.

Şahin, S. (2008). An application of peer assessment in higher education. *The Turkish Online Journal of Educational Technology, 7*(2), 5−10. ERIC Number: EJ1102895

Şahin, M. G., Taşdelen Teker, G. & Güler, N. (2016). An analysis of peer assessment through many facet Rasch model. *Journal of Education and Practice, 7*(32), 172−181. ERIC Number: EJ1122446

Topping, K. J. (1998). Peer assessment between students in colleges and university. *Review of Educational Research 68*(3), 249-276. https://doi.org/10.3102/00346543068003249

Topping, K. J. (2009). Peer assessment. *Theory into Practice, 48*, 20-27 https://doi.org/10.1080/00405840802577569

Tseng, S-C. & Tsai, C-C. (2007). On-line peer assessment and the role of the peer feedback: a study of high school computer course. *Computers and Education, 49*(4), 1161-1174. https://doi.org/10.1016/j.compedu.2006.01.007

Vygotsky, L. S. (1986). *Thought and Language*. MIT Press.

Wiliam, D. & Leahy, S. (2015). *Embedding Formative assessment: Practical Techniques for K-12 Classrooms.* LearningSciencesInternational.

Yin, R. K. (2007). *Case study research and applications. Design and methods*. Sage.